Individual Performance in the Beer Game:
Underweighting the supply line and the impact of personality.

Brad C. Meyer
Drake University

Radostina Purvanova
Drake University

Data was collected from undergraduate students playing an online version of the Beer Game and correlated with data from a personality survey. Several measures of individual performance were developed that loaded into 5 factors: anchoring, weighting of visible inventory, weighting of unseen inventory, inconsistency, and amplification. Results show that personality does predict some decision behavior in the beer game, and more so for the wholesaler and distributor roles.

Introduction

The Beer Distribution Game is a classical activity used in operations, supply chain management, and system dynamics education to demonstrate what is commonly called the "bullwhip effect," that is, demand amplification in a supply chain (Sterman 1989, 1992). This phenomenon is not merely a classroom artifact. Papers written as early as 100 years ago called attention to demand amplification (Geary, 2006) and empirical evidence of amplification in various companies and industries has been quantified and reported (Holmstrom, 1997, Metters, 1997, Blinder, 1986, Blanchard, 1983, West, 1986, Krane and Braun, 1991, Cachon, Randall, Schmidt, 2007).

Jay Forrester described the bullwhip effect as arising from the dynamics of a system with time delays and feedback loops (Forrester, 1958, 1961). He attributed the phenomenon to human behavior in its inability to understand the complexities of a system. Since the time of Forrester's pioneering work, numerous others have examined the impact of human behavior in supply chain systems and recommended various remedies or ameliorators of the bullwhip effect (Sterman, 1989, Croson and Donahue, 2006, Wu and Katok, 2006, Cantor and MacDonald, 2009, Nienhaus, 2006).

One of the authors of this paper has used the Beer Game in class for over 10 years making casual observations about performance in the game. Final total costs for the game vary among the teams in the class commonly by a factor of 5 to 10. That is, the worst total cost is 5 to 10 times the amount of the best total cost. Some students seem to greatly exaggerate the demand they receive, while others behave more conservatively. From an anecdotal perspective, it appears that when a team does exceptionally well, the students on the team are not just a random draw from the class. When a team does exceptionally bad, it is not uncommon to look at the members of the team and say, "that makes sense, given those students."

Anecdotal evidence proves nothing, but it is a useful source for theory generation. Such observations instigated a conjecture that some of the difference in performance may be tied to traits of individuals, and not merely to random, unknowable causes. We collected data on game decisions and personality for over 130 participants using an online version of the Beer Game in order to test several hypotheses regarding the impact of personality on behavior in the Beer Game.

We begin the paper with a brief overview of the Big Five theory of personality and research that relates personality to decision making. Next, we describe the Beer Game as used in this research and our experimental procedures. We describe and demonstrate our methodology to isolate individual behavior and performance in a system where the decisions made are clearly interdependent. Finally, we look at the results of our analysis to relate personality to game behavior. We close with conclusions and recommendations for further work.

Brief Introduction to the Big Five Personality Traits.

Though the study of personality has a long history dating back at least to the end of the nineteenth century when Galton first attempted to uncover the structure of personality (Galton, 1884), the importance of personality variables in applied settings did not become apparent until the end of the twentieth century. It took applied scientists a whole century to realize the usefulness of personality for predicting individual, team and organizational performance and effectiveness mainly because a commonly agreed-upon personality model was lacking. Today, most social scientists and applied behavior researchers embrace the Five-Factor Model (better known simply as The Big Five). Hundreds of studies, including cross-cultural studies, have successfully replicated the model, and have validated its predictive validity in a broad range of contexts (Goldberg, 1995).

The Big Five personality traits are Extraversion, Neuroticism, Conscientiousness, Agreeableness and Openness to Experience. Extraversion refers to individuals' tendencies to be sociable, dominant, and energetic. Neuroticism describes individuals' tendencies to exhibit low self-esteem, to experience negative emotions such as anger, self-doubt and depression, and to act in moody and temperamental ways. Extraversion and Neuroticism are the most stable out of the Big Five; in other words, they are the two traits that are replicated nearly without fail across contexts. Conscientiousness is also an easily replicable trait; it captures individuals' tendencies to be dutiful and responsible, achievement-oriented and internally motivated, and cautious, risk-averse and control-oriented. Importantly, Conscientiousness is the best predictor of individual job performance across occupations (followed closely by Neuroticism). The last two Big Five traits—Agreeableness and Openness to Experience—are less robust, as their theoretical structure is less often replicated in consistent ways. Still, these traits have proven useful for predicting relational skills (Agreeableness) and performance on intellectual tasks (Openness). Specifically, Agreeableness captures individuals' propensity to be nurturing, sympathetic and caring, to cooperate and act in altruistic ways, as well as to trust others and be fair to others. Openness to Experience describes individuals' curiosity and creativity, as well as their intellect and intellectual complexity.

In the developing field of Behavioral Operations Management, few studies have examined the role of personality in the Beer Game; in fact, we are aware of only one such study—Ruel, van Donk and van der Vaart (2006). These authors focused on the impact of four personality traits—risk-taking orientation, tolerance for ambiguity, self-efficacy, and locus of control—on decision-making in the Beer Game. They found that low risk-takers have higher backorder costs and lower inventory costs than higher risk-takers. Ruel et al. also attempted to investigate whether risk-taking orientation interacts with participants' role in the Beer Game (for example, does risk-taking affect behavior differently for retailers vs. wholesalers), but due to their small sample size, their results were inconclusive. We expand on this research by looking at a broader range of personality traits in a larger sample. Further, while these researchers formed supply chain teams composed of students with similar personality, we did not manipulate the personality composition of teams, choosing instead to obverse the effects of personality in a more natural setting.

When Is Personality More Relevant?

Though the role of personality for predicting human behavior is well documented and accepted, some have argued that it might have been at least somewhat overplayed. At issue is the idea that not all situations are created equal, and that personality may be a more or less important predictor of behavior under some conditions than under others. Mischel's (1977) famous distinction between "strong" and "weak" situations has proven useful in explaining this notion. Strong situations are structured and clear. This means that strong situations lead everyone to interpret the situation the same way, as well as to behave in a predictable and commonly agreed-upon manner. In contrast, weak situations are ambiguous and ill-defined; they are filled with uncertainty and unpredictability. Importantly, this means that weak situations do not provide cues as to what constitutes appropriate behavior. Mischel argues, and much research since has documented (Adams, Roch, & Ayman, 2005; Purvanova & Bono, 2009; Shamir & Howell, 1999; Waldman & Yammarino, 1999) that personality tends to have a greater effect on behavior in weak than in strong situations.

In the context of the Beer Game, ambiguity and uncertainty arise from two sources: demand from the [direct] customer and stock shipments from the supplier. However, as the game is construed, the amount of uncertainty is not the same for all members of the chain. The demand stream to the retailer is almost constant. Any uncertainty occurs early in the game when demand doubles, after that the demand signals to the retailer constitute a stronger situation than for the other members in the chain. The factory also has reduced ambiguity. Whatever amount the factory orders will arrive 3 weeks later. There is never any uncertainty in the supply side for the factory. In contrast, the situations for the wholesaler and distributor are filled with uncertainty compared to those of the retailer and factory, and therefore constitute "weak" situations. This implies that the impact of personality will be greater for the wholesaler and distributor positions compared to the retailer and factory positions.

The Beer Game and Experiment Procedures

The classic Beer Game, as designed by the MIT Systems Dynamics group in the 60s, assigns players to simple, four member supply chains. Each supply chain consists of a retailer, wholesaler, distributor, and factory. End customer demand is presented to the retailer who fills orders to the customer and issues replenishment requests from the wholesaler. The wholesaler orders stock from the distributor and the distributor from the factory. The factory has no supplier, but replenishes stock by production. The lead-time for production at the factory is a constant 3 weeks with no maximum quantity restrictions placed on factory production. An order to a supplier is not seen by the supplier until the second week after the order is placed. When an order is shipped to the immediate customer, it passes through two shipping delays such that it appears the second week after it leaves the supplier. Combined, this makes a four week lead time for replenishment with a potential for more weeks if the supplier is out of stock. Back orders are handled in such a way that no order is ever forfeited. Every order received will eventually be filled once the supplier has stock on hand.

The game is played on a game "board" (or multiple boards, one for each supply chain member) and plastic chips were used to represent cases of beer. Orders were written on slips of paper that were passed from customer to supplier. Cases of beer (chips) were passed from supplier to customer, spending a simulated week in each of two shipping delays. The player (or players, as often two persons were paired to operate one supply chain member) could easily see their own stock on hand and the amount in route, visible in the two shipping delays. When the game is played in a physical space, it is also possible for the players to glance down the table and see stock in the rest of the chain, but the players are generally briefed to ignore what is happening in the rest of the chain and not to talk or exchange information with the other members of the supply chain, other than orders and shipments.

Several computerized versions of the Beer Game have been developed (Kaminsky and Simchi-Levi 1998, Jacobs 2000, Kalidindi 2001). The online game used in this study, named Beer Chain, was developed by one of the authors and modeled after the classic Beer Game. The graphics were designed to provide a similar psychological effect that the chips provided in the physical game. Instead of merely displaying a numerical stock count, the game draws the stack of beer cases in the member's inventory and in the shipping delays. (Backlog, however, is displayed as a number.) See figures 1 and 2.

The Beer Chain software computes inventory on hand and backlog for each period as well as holding and backlog cost for each period. All decisions made by the members are stored in a database that can be queried after the game. The players are instructed that the goal of the exercise is to minimize costs. As in the classic Beer Game, the two costs incurred at the end of each simulated week are inventory cost, assessed at a rate of $0.5 per item in inventory, and backlog cost, assessed at a rate of $1 per item promised but not yet shipped.

For our experiments, undergraduate students in an introductory class in Operations Management were instructed to bring laptops to class and were assigned to chains and roles that were randomly distributed throughout the room, so that each player was not able to discern the other members of his or her chain. We implemented the typical demand stream of 4 cases for each of the first four weeks, increased to 8 cases in week 5 and remaining at 8 cases for the duration of the game.

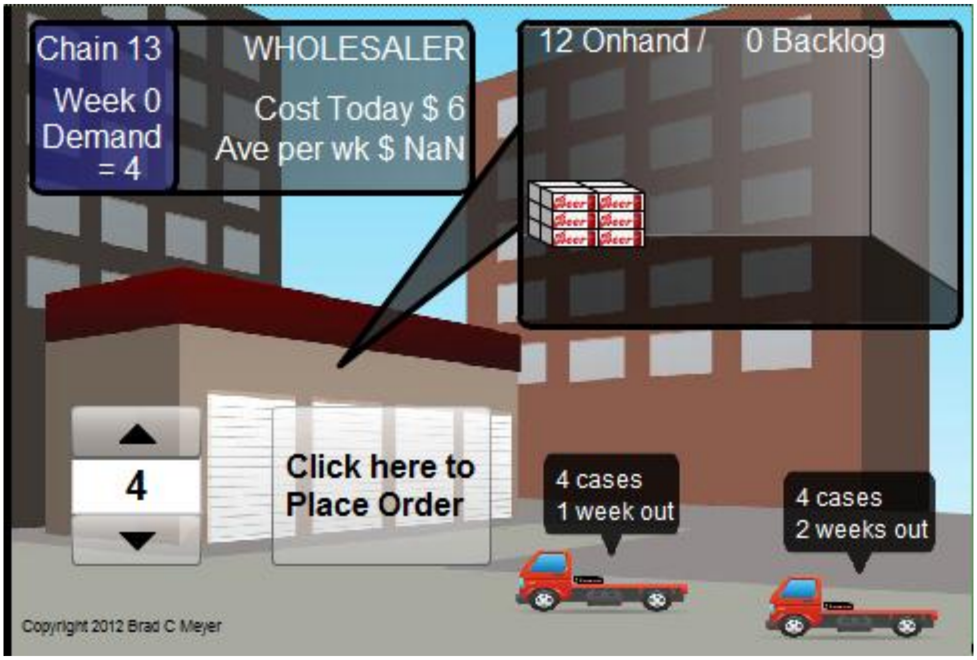

Figure 1 – Beer Chain graphics for the wholesaler role, in the first week of a game.

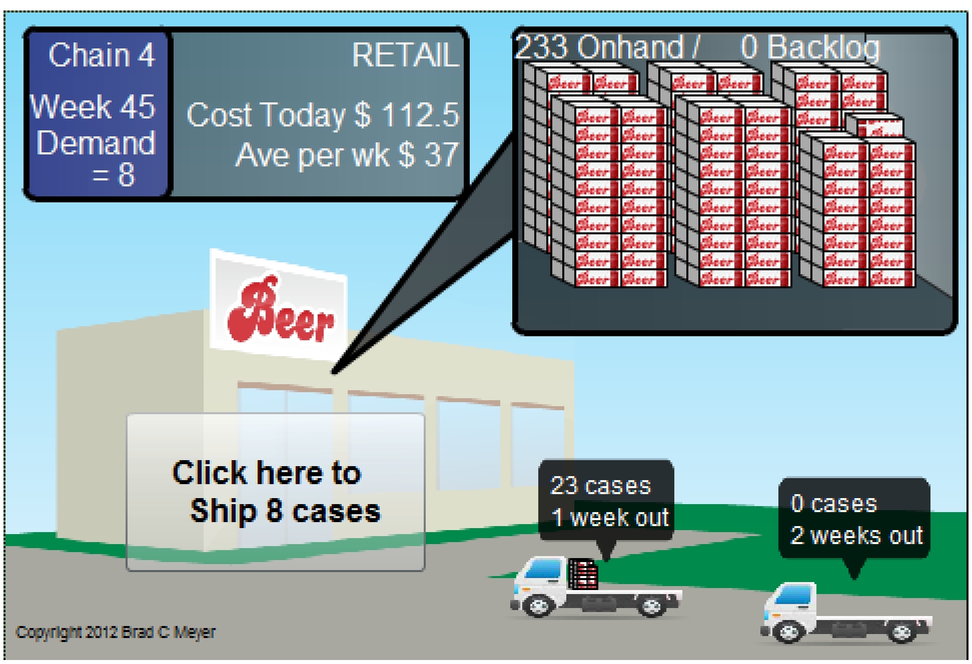

Figure 2 – Beer Chain graphics for the retailer role, at a later week in a game.

Goldberg's (1992) 50-item bipolar transparent inventory was used to measure the Big Five personality traits. This inventory consists of five sets of 10 bipolar adjective pairs—one set for each Big Five trait. For example, Extraversion was assessed with adjective pairs such as silent—talkative and unsociable—sociable, Agreeableness was assessed with adjective pairs such as cold—warm and distrustful—trustful, Conscientiousness was assessed with adjective pairs such as disorganized—organized and rash—cautious, Emotional Stability was assessed with adjective pairs such as angry—calm and unstable—stable, and Openness to Experience was assessed with adjective pairs such as imperceptive—perceptive and uncreative—creative. Adjective pairs were presented in two columns; column A listed all adjectives consistent with low levels of Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience, whereas column B listed the bipolar opposites of the adjective (i.e., adjectives consistent with high levels of Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience). The 1-to-9 rating scale was placed in-between the adjectives in each adjective pair, and respondents were asked to describe themselves as being rated (1) *very* (Trait A) through (5) *neither* (Trait A) *nor* (Trait B) to (9) *very* (Trait B). The scores on each set of 10 bipolar adjectives were averaged to form five overall trait ratings, one for each Big Five trait, with a high score indicating a high level of that trait.

Determining individual member performance

Performance of a chain is measured by the sum of inventory and backlog costs over all 4 members of the supply chain over the duration of the game. The best performance with a simple ordering rule, "order exactly what was demanded" yields a total cost of $414 for a 48 week game. Typical results for actual players for the Beer Chain version of the game, for a 48 week duration, range from $2,500 to $15,000, indicating that demand amplification is clearly present. Since the chain is interdependent, the costs at one position cannot simply be taken as a measure of the performance of the person playing that position.

The inventory cost at a position can be thought of as being comprised of two components. One is the cost of safety stock. A member may choose (directly or indirectly) to carry some stock to guard against a larger than average order or to protect in case the supplier is not able to deliver in a timely way. Safety stock at a position is determined by the player, but that determination is based upon an assessment of uncertainty in customer demand and uncertainty in lead time from the supplier. If the customer is sending inconsistent demand signals, or if the supplier is not fulfilling requests in a timely fashion, this could cause the member to increase the amount of safety stock carried. So, while it derives from the member's decision, the member must react to the signals received which could result from poor performance either upstream or downstream.

The other component of inventory cost in this game derives from over-ordering. If a member orders more than is being requested, eventually that stock will appear in the member's inventory and remain until called out by the customer. Two potential causes of over-ordering are failure to account for inventory in the supply line and extrapolating demand when incoming orders are increasing. The underweighting of supply line inventory is an error, and could be identified as poor performance. Extrapolating demand, on the other hand, may be wise behavior in general, but if caused by bad demand signals, it will yield higher costs for the chain.

Note that over-ordering will not add to downstream holding costs and the over-ordered stock itself will not add to upstream holding cost. This is because inventory cost is only assessed on end-of-period inventory on hand. Over-ordered stock will pass through upstream suppliers from shipping delay to shipped amount in the same period and never incur holding cost. The only way over-ordering will increase upstream holding costs is if the over-ordering occurs in such a way as to cause the upstream members to further amplify demand.

In summary, holding cost at a position is largely determined by the member at that position, but also influenced by signals of customer orders and supplier lead times.

Backlog cost, on the other hand, is primarily the fault of the position's upstream or downstream partners. Backlog occurs when a member is not able to fulfill demand. Inability to fill demand occurs when the customer asks for an amount that is not in line with the signals that have been sent previously. In particular, when the customer amplifies demand beyond what the member can cover with safety stock. It is exacerbated when the upstream chain also cannot handle the increased demand level, and the replenishment lead time lengthens. A member will cause her own backlog when she is not carrying the amount of safety stock appropriate to the signals being sent so far. When there is an increase in end customer demand, there will of necessity be backlog until the chain adjusts to the new demand level. But unnecessary backlog is caused by a customer who is over-ordering, and less often by a supplier who is under-ordering.

To assess performance of individual members in the supply chain it is necessary to determine the degree to which a player is under-ordering and over-ordering with respect to the signals she is receiving. We now explain a set of such measures which can be culled from game data. These are classified into three types:

1. Measures of demand amplification
2. Measures of departure from "ideal" decisions
3. Measures based on regression of players game models.

1. Measures of demand amplification

The average demand can be computed for the end customer (demand of 4,4,4,4 then 8 for the rest of the game) and each of the 4 members, from their order records. As has been shown in previous studies, the demand tends to increase upstream in the chain. For example, table 2 shows the average demand for each position for 15 teams who played in spring of 2013. With some exceptions, average demand increases from customer to factory.

Team	Customer	Retailer	Wholesaler	Distributor	Factory
1	7.67	7.54	15.73	17.29	22.63
2	7.67	7.75	7.58	8.52	9.52
3	7.67	8.02	7.77	7.75	11.44
4	7.67	7.79	9.00	8.98	9.08
5	7.67	8.79	9.83	10.98	12.29
6	7.67	7.77	12.75	21.98	26.83
7	7.67	7.96	12.21	12.65	13.73
8	7.67	7.67	7.33	11.38	16.17
9	7.67	7.50	7.96	11.29	17.63
10	7.67	7.90	8.15	11.92	14.31
11	7.67	9.67	14.17	19.02	19.88
12	7.67	7.75	9.21	21.42	27.35
13	7.67	8.54	7.79	7.54	14.81
14	7.67	7.65	8.71	14.17	23.06
15	7.67	7.75	8.19	16.98	11.56
Average	**7.67**	**8.00**	**9.76**	**13.46**	**16.69**

Table 2 – Average demand for each position of the supply chain for 15 teams playing spring of 2013.

In order to find a demand amplification measure that shows the degree to which an individual member amplified demand, we divide each average demand by the member's customer's demand. Averaging for each role, we find that the average amplification for the retailer is 1.04, for the wholesaler, 1.22, for the distributor, 1.39 and for the factory, 1.27.

Similar tables can be computed for standard deviation of demand for each member and for the amplification of standard deviation. The average values differ by position, with statistical significance at the .05 level, using data from 2 semesters. Therefore we normalized the amplification by dividing by the average for that position giving two individual performance measures in this category.

a. normalized demand amplification
b. normalized variation amplification

2. Measures of departure from idealized performance

A simple, but near optimal decision strategy is to always order what your customer orders. This may not be best for all demand streams, but for this version of the game it gives an excellent result. So we develop a measure of departure from this ideal strategy based on the absolute difference between the amount demanded in a period and the amount ordered. Table 3 shows the average absolute differences for the same 15 chains as before.

Chain	1	2	3	4
1	6.54	17.81	15.90	13.71
2	3.08	4.25	3.31	7.63
3	0.98	2.79	0.29	9.90
4	0.75	4.33	3.96	3.94
5	2.79	4.17	5.06	4.77
6	2.60	10.10	18.52	10.60
7	4.96	7.04	0.98	10.04
8	5.75	2.21	14.02	7.46
9	7.00	4.63	7.00	11.25
10	1.31	4.04	7.15	6.15
11	6.92	12.75	12.10	14.90
12	6.54	5.92	16.46	25.52
13	3.04	0.42	5.75	9.77
14	2.56	8.63	8.75	9.56
15	3.00	4.73	9.42	12.54
Average	**3.86**	**6.25**	**8.58**	**10.52**

Table 3 – Average absolute difference between amount ordered and customer demand

Clearly, the value grows upstream in the chain as the average demand grows upstream in the chain. In this case, we can divide each value by the average demand for that position to get a more comparable measure as shown in Table 4. The average values are much closer, suggesting that dividing by demand has corrected for the growing demand across the positions. But it may not have corrected for the difference in uncertainty or other factors across the chain, so we also normalize the value by position average, yielding one measure for this category.

> *c. normalized average absolute difference between amount ordered and amount demanded divided by average demand.*

3. Measures based on regression of player's game models.

The first two categories of measures seek to assess individual performance by using the decision values (orders) directly. Sterman (1989) and Croson and Donahue (2006) distinguish individual behavior by developing a rational model of the decision maker's behavior and using decision data to determine the parameters of the model. Sterman's model incorporated a demand forecast based on past demand and a tendency to anchor. The Croson and Donahue model was somewhat simpler. In both cases, they modeled the human as having the following signals: incoming order, incoming shipment, inventory in the supply line, and inventory on hand (positive or negative). Sterman also incorporated the pattern of past demands as a signal used to forecast future demand.

Chain	1	2	3	4
1	0.87	1.13	0.92	0.61
2	0.40	0.56	0.39	0.80
3	0.12	0.36	0.04	0.87
4	0.10	0.48	0.44	0.43
5	0.32	0.42	0.46	0.39
6	0.34	0.79	0.84	0.40
7	0.62	0.58	0.08	0.73
8	0.75	0.30	1.23	0.46
9	0.93	0.58	0.62	0.64
10	0.17	0.50	0.60	0.43
11	0.72	0.90	0.64	0.75
12	0.84	0.64	0.77	0.93
13	0.36	0.05	0.76	0.66
14	0.34	0.99	0.62	0.41
15	0.39	0.58	0.55	1.08
Average	**0.48**	**0.59**	**0.60**	**0.64**

Table 4 – Average absolute difference divided by average demand

After testing several options, we used the Croson and Donahue model with small modifications. Equation 6 in their paper describes the regression model as

$$O_t^{ig} = \max\{0, \alpha_0 + \alpha_I I_{t-1}^{ig} + \alpha_R R_t^{ig} + \alpha_S S_t^{ig} + \alpha_N N_t^{ig} + \alpha_t t + \varepsilon\}$$

where *i* is the role number, *g* is the team number, and *t* is the time period. O signifies amount ordered, I the on hand inventory level, positive or negative, R the order received from the immediate customer, S the shipment received from the immediate supplier, and N the member's total outstanding orders in the period.

We considered the particulars of our version of the game and reasoned that since the decision makers were able, at the time of the decision, to see the inventory or backlog on hand and the amount in the two shipping delays, but could no longer see what had been supplied to them that week, we replaced the shipments received and outstanding orders variables with two supply line variables: one for the amount visible in the supply line (in the two shipping delays) and one for inventory ordered and not yet visible in the supply line. Croson and Donahue included t to test for any effects of time. After testing, we removed t to simplify the model.

Letting V represent the amount visible in the supply line and U the amount unseen in the supply line, our revised model was

$$O_t^{ig} = \max\{0, \alpha_0 + \alpha_I I_{t-1}^{ig} + \alpha_R R_t^{ig} + \alpha_V V_t^{ig} + \alpha_U U_t^{ig} + \varepsilon\}$$

Figure 3 shows a causal diagram, patterned after Sterman(2000), that represents the situation.

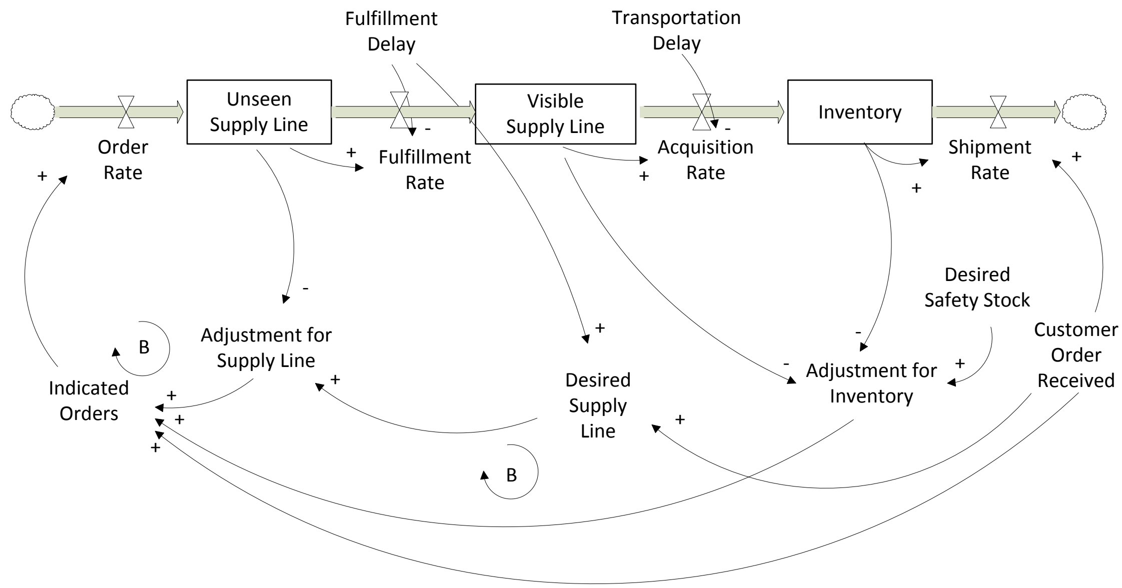

Figure 3 – Causal Diagram for a single member of the Beer Game supply chain.

Using censored regression in R, we computed the best fit alphas for each participant. Thus the following performance measures were generated:

d. member's safety stock coefficient (α_0)
e. member's coefficient on inventory on hand (α_I)
f. member's coefficient on incoming demand (order received) (α_R)
g. member's coefficient on inventory visible in the supply line (α_V)
h. member's coefficient on inventory unseen in the supply line (α_U)
i. the RSQ value for the member's model fit to the decisions

Next we performed a factor analysis on these 9 factors, the results of which are shown in table 5. Four factors emerge from the analysis.

Factor 1 Anchoring. The safety stock level and the weighting of demand are highly but inversely related. The more weight put on demand, the lower the safety stock coefficient. We labeled this a form of anchoring, with the anchor being the desired safety stock level. Those with high anchoring based their order on a desired safety stock level, putting little weight on current demand. Those with low anchoring based their order on the amount demanded, putting little weight on maintaining a fixed safety stock level.

Rotated Component Matrix

	Component			
	1	2	3	4
Safety Stock	-.904	-.308	.027	-.019
Inventory Onhand	.209	.781	-.159	-.313
Incoming Demand	.952	-.139	-.063	.055
Supply Line Visible	.074	.768	.012	.004
Supply Line Unseen	-.252	.617	.274	.320
RSQ	-.087	-.032	.943	-.111
Normalized Demand Amplification	.001	.048	.043	.839
Normalized Variation Amplification	.135	-.020	-.100	.894
Normalized Average Absolute Difference	-.104	-.181	-.451	.635

Table 5 Component matrix.

Factor 2 Weighting of Inventory. The weight put on inventory, whether visible on hand, visible in the supply line, or unseen in the supply line, emerged as a factor. This is not to say that the weights on these three are equal, but that they tend to be higher or lower as a set, not behaving independently. Although these components load together, we split this into two factors: weighting of visible inventory (2a) and weighting of unseen inventory (2b). This highlights the underweighting of unseen inventory which is generally understood to be a cause of demand amplification.

Factor 3 Inconsistency. This factor consists only of the RSQ value for the regressions. A larger value means the theoretical model fit the participant's behavior well. A low RSQ means that the best fit model was not a good predictor of the participant's decisions, perhaps because the participant changed her mental model as the game progressed, or simply due to irrational behavior, (with the possibility that the participant used a rational model that was substantially different from the one we assumed.)

Factor 4 Amplification. The two measures of amplification, that of demand and of variation of demand, load together. We find that our measure of difference from ideal loads with them, indicating that it also is, in effect, a measure of the tendency to amplify demand.

Predictions

Based on our theoretical analysis of behavior in the Beer Game, we expect the following to be supported by our results.

1. Weight coefficients for onhand inventory and unseen supply line inventory will be comparable to those found by Sterman (1989) and Croson and Donahue (2006).

2. The underweighting of unseen inventory will correlate with amplification, since underweighting of inventory is generally considered to be a cause of demand amplification.

3. Extraversion will predict the underweighting effect and demand amplification. Extraverted individuals are typically more optimistic; that is, they more often expect to receive positive rather than negative outcomes, and they experience positive emotions and are in a positive mood more often. Although optimism and positive moods are generally linked to well-being and other positive psychosocial outcomes, they can also lead to unimpressive task performance. This is because individuals in a positive mood are more likely than individuals in negative mood to perceive that they have achieved their goals or made sufficient or "good enough" progress on tasks, and as a result, they are more likely to stop exerting effort (George & Zhou, 2002).

4. Conscientiousness will predict underweighting; specifically, high conscientiousness will induce less underweighting. Conscientious individuals are usually internally motivated and approach their tasks with a great degree of responsibility. These qualities might help Conscientious individuals overcome the underweighting effect by paying more careful attention to when and how orders are placed.

5. Openness will predict the underweighting factor. Individuals who are open to experience typically tolerate, and even seek out ambiguity; they feel comfortable in underdefined, unstructured contexts. Such tolerance for ambiguity is a handy asset as it allows for more cognitive resources to be freed to handle ambiguity more seamlessly and successfully. This suggests that open individuals will be less often subject to the underweighting bias than less open individuals. Perhaps more importantly, open individuals not only tolerate ambiguity better, but their intellect and cognitive complexity allows them to face challenges more successfully (which is why they might tolerate ambiguity so well in the first place). Open individuals are intelligent, perceptive and analytical, as well as creative and imaginative. These are all traits which should help Open individuals overcome the underweighting effect by utilizing their superior cognitive resources to their advantage. If they are more complex thinkers, they will have loose bounds on their rationality.

6. Personality effects will be more pronounced for the wholesaler and distributor than for the retailer and factory, due to differing amounts of ambiguity in their roles.

Results

1. In comparison to the two above mentioned studies, we found similar average weighting coefficients. This again shows the tendency for players to underweight the

supply line inventory. In our study, we separated the weighting of the seen and unseen supply line inventory. In order to compare to earlier studies, we also ran the individual member regressions with the two combined, yielding the value in the table 6 below.

Study	average weight on on-hand inventory	average weight on supply line inventory
Sterman	-0.26	-0.0884
Croson and Donahue	-0.2368	-0.0302
This study	-0.2653	-0.0679

Table 6 Comparison of coefficients to previous studies.

2. In our factor analysis (see table 5 above), the indicator of underweighting did not load on the same factor as the three indicators of demand amplification. Further, as Table 7 below shows, the factor called "weighting of unseen inventory" (i.e., the underweighting effect), was only marginally correlated with the factor called "amplification". However, when we look at only those who played the role of distributor or wholesaler, we find that the relationship between amplification and weighting of unseen inventory is strongly significant ($p < .01$), as predicted. Unseen inventory plays a very minor role for the factory, which may be washing out the strength of the effect in the entire data set.

Our next set of predictions involved relationships between the Big Five personality traits, and underweighting and amplification. Results are displayed in Tables 7, 8, and 9 below.

3. We find that across all four roles (Table 7 below), extroversion predicts amplification, as expected. Specifically, individuals with higher levels of Extraversion are more likely to amplify demand than individuals with lower levels of Extraversion ($r = .19^{*}$). Contrary to expectations, Extraversion did not predict underweighting of the unseen inventory. We also observed a marginally significant association between Extraversion and Inconsistency ($r = -.16^{\dagger}$). Though not predicted, we believe this finding fits with our general theory of the role Extraversion plays in the context of the Beer Game. Extraverted individuals are more impulsive, as we previously discussed, and this might explain why they are less likely to stick to one rational mental model and to play the game in an inconsistent manner.

4. We find that across all four roles (Table 7), conscientiousness does not have any significant impact on behavior.

5. We find that across all four roles (Table 7), openness does not predict underweighting, though as predicted, it is marginally related to amplification ($r = .14^{\dagger}$). Interestingly, these effects run counter to what we had predicted. We hypothesized that Open individuals, because of their intellect and interest in thinking, would be less likely to underweight and amplify. However, we actually observed that Open individuals are marginally more likely to amplify than less Open individuals. We also saw that they

were less likely to employ a consistent, rational mental model. We are presently investigating the reasons for these unexpected findings.

6. Our sixth and last prediction expected that the effects of personality would be especially pronounced under conditions of greater situational uncertainty, such as that experienced by the players occupying the middle two roles (wholesaler and distributor). Table 8 shows the correlations we observed within the sample of retailers and factory only (lower situational uncertainty), whereas Table 9 shows the correlations we observed with the sample of wholesalers and distributors only (higher situational uncertainty). As seen in Table 8, all the effects of personality we had previously observed disappeared in the lower uncertainty condition. In contrast, and as predicted, Table 9 shows that the pattern of correlations we had previously observed holds among the wholesalers and distributors, i.e., in the higher uncertainty condition. Perhaps the most important observation in Table 9 is that some of the predicted associations between personality and weighting of unseen inventory emerge here. Specifically, as predicted in prediction #3, Extraverted individuals have higher values on underweighting unseen inventory ($r = .20^{\dagger}$). However, contrary to prediction #5, Open individuals also have higher values on underweighting unseen inventory ($r = .28^{*}$). Furthermore, we also observe that Agreeable individuals in roles of higher uncertainty place higher values on underweighting unseen inventory ($r = .21^{\dagger}$). We are currently working on clarifying the meaning of the observed associations.

	1	2	3	4	5	6	7	8	9	10
1. Extraversion	--									
2. Conscientiousness	.02									
3. Openness	.38**	.29**								
4. Agreeableness	.26**	.24**	.33**							
5. Neuroticism	.33**	.13	.29**	.28**						
6. Anchoring	-.04	.00	-.05	-.02	-.12					
7. Weighting of visible inventory	.03	.01	.08	.14	$.15^{\dagger}$	-.44**				
8. Weighting of unseen inventory	-.07	.00	.08	.12	.12	-.07	.26**			
9. Inconsistency	$-.16^{\dagger}$	.00	-.17*	.12	-.02	-.01	-.03	.17*		
10. Amplification	.19*	-.02	$.14^{\dagger}$	-.04	.04	-.06	-.14	$.15^{\dagger}$	-.26**	--

Table 7 Zero-Order Correlations Among Study Variables, Entire Dataset
N=133; **p<.01, *p<.05, †p<.10.

	1	2	3	4	5	6	7	8	9	10
1. Extraversion	--									
2. Conscientiousness	-.08									
3. Openness	.29*	.23†								
4. Agreeableness	.29*	.06	.27*							
5. Neuroticism	.38**	.09	.28*	.32**						
6. Anchoring	-.09	.05	.07	-.07	-.15					
7. Weighting of visible inventory	.09	-.01	-.14	.12	.13	-.39**				
8. Weighting of unseen inventory	-.14	-.03	-.10	.11	.11	-.02	.21†			
9. Inconsistency	-.19	-.05	-.20	.10	.05	.16	-.09	.31*		
10. Amplification	.18	.02	.08	.08	.00	.08	-.07	.03	-.21†	--

Table 8: Zero-Order Correlations among Study Variables, Retailer and Factory Only N=66; **p<.01, *p<.05, †p<.10.

	1	2	3	4	5	6	7	8	9	10
1. Extraversion	--									
2. Conscientiousness	.05									
3. Openness	.45**	.27*								
4. Agreeableness	.09	.30*	.30*							
5. Neuroticism	.23*	.17	.30*	.23*						
6. Anchoring	-.01	-.03	-.11	.01	-.11					
7. Weighting of visible inventory	.04	.01	.19	.18	.17	-.46**				
8. Weighting of unseen inventory	.20†	.02	.28*	.21†	.15	-.12	.33**			
9. Inconsistency	-.19	.08	-.14	.13	-.11	-.13	.03	.04		
10. Amplification	.22*	-.06	.19	-.14	.08	-.14	-.18	.33**	-.32**	--

Table 9: Zero-Order Correlations among Study Variables, Wholesaler and Distributor Only. N=67; **p<.01, *p<.05, †p<.10.

Summary and Suggestions for Further Study

Our results are consistent with previous work in demonstrating the phenomenon of underweighting of the supply line. We were also able to show correlation of underweighting the supply chain with demand amplification. This was marginally significant overall and strongly significant for the wholesaler and distributor roles.

The role of personality in the Beer game was shown to be dampened for roles with less uncertainty, that is, the factory and the retailer roles in this study. Extroversion predicted demand amplification and thus higher costs. Extroversion predicted underweighting of supply line inventory, but only for distributors and wholesalers, and the relationship was only significant at the p<.10 level. Perhaps extroversion induces amplification in other ways, such as inconsistency and by inducing extrapolation of demand growth. This can be explored in future work.

The influence of conscientiousness and openness was not apparent or differed from our expectations. There may be other ways of exploring this. For example, the trait of openness can be divided into sub-traits. It may be that one or more of the sub-traits does predict Beer Game decisions even if the aggregated trait does not.

The factor we labeled as anchoring needs more exploration. The inverse relationship between weight on demand and the safety stock level suggests that there may be more than one strategy used by participants in this game. Some may ignore inventory amounts and simply pass on orders received to their suppliers. Others may try to actively manage inventory amounts, with concerns for safety stock and supply line inventory. Perhaps there are repercussions for one strategy that are different than for the other. This is something else that needs further study.

References

Adams, S. J., Roch, S. G., & Ayman, R. (2005). Communication medium and member familiarity: The effects on decision time, accuracy, and satisfaction. *Small Group Research, 36,* 321-353.

Blanchard, O. J. (1983) The production and inventory behavior of the American automobile industry, J. *Polit. Econ., 91,* 3: 365-400.

Blinder,A.S., (1986) Can the production smoothing model of inventory behavior be saved? Q. J. *Econ., 101,* 3: 431-454.

Cachon, G. P., Randall, T, Schmidt, G. M. (2007) *Manufacturing & Service Operations Management; In Search of the Bullwhip Effect*, Fall 2007; 9, 4; ABI/INFORM Global pg. 457

Cantor, D. E., & Macdonald, J. R. (2009). Decision-making in the supply chain: examining problem solving approaches and information availability. *Journal of Operations Management*, *27*(3), 220-232.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48,* 306-307.

Carnevale, J. J., Inbar, Y., & Lerner, J. S. (2011). Individual differences in need for cognition and decision-making competence among leaders. *Personality and Individual Differences*, *51,* 274-278.

Croson, R., & Donohue, K. (2006). Behavioral Causes of the Bullwhip Effect and the Observed Value of Inventory Information. *Management Science*, (3), 323. doi:10.2307/20110512

Forrester, J. 1958. Industrial dynamics: A major breakthrough for decision makers. Harvard Bus. Rev. 36 37-66.

Forrester, J. W., *Industrial Dynamics,* MIT Press, Cambridge, MA, 1961.

Galton, F. (1884). Measurement of character. *Fortnightly Review, 36,* 179-185.

Geary, S., Disney, S.M., and Towill, D.R., 2006. On bullwhip in supply chains – historical review, present practice and expected future impact. International Journal of Production Economics, 101 (1), 2–18.

George, J. M., & Zhou, J. (2002). Understanding when bad moods foster creativity and good ones don’t: The role of context and clarity of feelings. *Journal of Applied Psychology, 87,* 687-697.

Goldberg, L. R. (1995). What the hell took so long? Donald W. Fiske and the Big-Five factor structure. In P. E. Shrout & S. T. Fiske (Eds.). *Personality research, methods and theory: A festschrift honoring Donald W. Fiske* (pp. 29-43). Hillsdale, NJ: Lawrence Erlbaum Associates.

Holmstro¨ m, J., 1997. Product range management: A case study of supply chain operations in the European grocery industry. International Journal of Supply Chain Management 2 (8), 107–115.

Hough, L. M., & Ones, D. S. (2002). The structure, measurement, validity, and use of personality variables in industrial and organizational psychology. In N. Anderson, D. S. Ones, H. K. Sinangil & C. Viswesvaran (Eds.). *handbook of Industrial, Work, and Organizational Psychology, V. 1* (pp. 233-277). Thousand Oaks, CA: Sage Publications.

Hough, L. M., & Oswald, F. L. (2000). Personnel selection: Looking toward the future – remembering the past. *Annual Review of Psychology, 51,* 631-664.

Jacobs, F.R., 2000. Playing the Beer Distribution Game over the Internet, *Production and Operations Management* 9 (1) 31-39.

Kalidindi, V. S. 2001. The web based beer distribution game: An application. Unpublished Masters thesis, Department of Mechanical Engineering, University of Minnesota, Minneapolis, MN.

Kaminsky, P., D. Simchi-Levi. 1998. A new computerized beer distribution game: Teaching the value of integrated supply chain management. H. Lee, S. M. Ng, eds. Global Supply Chain andTechnology Management, Vol. 1. The Production and Operations Management Society Series in Technology

Krane, S.D. and Braun, S. N.. Production smoothing evidence from physical product data, J. *Polit. Econ., 99,* 3 (1991) 558-581.

Manor, J. K., Richey, J. A., Cromer, K., Mallott, M., Lejuez, C. W., Joiner, T. E, & Schmidt, N. B. (2007). Dispositional anxiety and risk-avoidant decision-making. *Personality and Individual Differences, 42*, 665-675.

McAdams, D. P. (1992). The five-factor model *in* personality: A critical appraisal. *Journal of Personality, 60*, 329-361.

Metters, R., (1997). Quantifying the bullwhip effect in supply chains. Journal of Operations Management, 15 (2), 89–100.

Mischel, W. (1977). The interaction of person and situation. In D. Magnusson & N. S. Endler (Eds.). *Personality at the crossroads: Current issues in interactional psychology.* Hillsdale, NJ: Erlbaum.

Nienhaus, J. How human behaviour amplifies the bullwhip effect. A study based on the beer distribution game online. *Production Planning & Control* Volume: 17 Issue: 6 (2006-09-01) p. 547-556. ISSN: 0953-7287

Purvanova, R. K., & Bono, J. E. (2009). Transformational leadership in context: Face-to-face and virtual teams. *Leadership Quarterly, 20,* 343-357

Ruël, G., van Donk, D. P., & van der Vaart, T. (2006). The beer game revisited: Relating risktaking behaviour and bullwhip effect. *Procedings of EurOMA*, 403-412.

Shamir, B., & Howell, J. M. (1999). Organizational and contextual influences on the emergence and effectiveness of charismatic leadership. *Leadership Quarterly, 10,* 257 – 283.

Sterman, J. D. (1989) Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment, *Manage. Sci., 35*, 3: 321-339.

Sterman, J.D., (2000) *Business Dynamics: Systems Thinking for a Complex World,* McGraw Hill.

Sterman, J.D., (1992) Teaching takes off, flight simulators for management education, *OR~MS Today, 19,* 5, 40-44.

Waldman, D. A., & Yammarino, F. J. (1999). CEO charismatic leadership: levels-of-management and levels-of-analysis effects. *Academy of Management Review, 24,* 266-285.

West, K. D., A variance bounds test of the linear quadratic inventory model, J. *Polit. Econ., 94,* 4 (1986) 374-401.

Wu, D. Y., & Katok, E. (2006). Learning, communication, and the bullwhip effect. *Journal of Operations Management, 24*(6), 839-850.