

Ludwig, T.D., Biggs, J., Wagner, S., & Geller, E.S. (2002). Using public feedback and competitive rewards to increase the Safe Driving Behaviors of Pizza Deliverers. *Journal of Organizational Behavior Management*, 21 (4), 75-104. Version of record published by Taylor & Francis and is available online at: <http://www.informaworld.com/> (ISSN: 0160-8061) DOI: 10.1300/J075v21n04_06 [Special Issue of the *Journal of Organizational Behavior Management* critiquing Ludwig (2002) included commentaries by Houchins & Boyce (2002), Austin & Wilson (2002), and Geller (2002).]

Keywords: Feedback | public posting | competitive rewards | safe driving | pizza deliveries, occupational safety | behavioral covariance | unobtrusive observation

Using Public Feedback and Competitive Rewards to Increase the Safe Driving of Pizza Deliverers

Timothy D. Ludwig, Jay Biggs, Sandra Wagner, & E. Scott Geller

ABSTRACT

This study investigated the effects of a safe driving competition facilitated by publicly-posted individual feedback. The turn-signal use, safety-belt use, and complete intersection stopping of 82 pizza deliverers were observed at two experimental and three control stores, as well as among the civilian population. After baseline observations, pizza deliverers received posted weekly individual feedback on their turn-signal use (at Store A) or complete intersection stops (at Store B). The deliverers' safe driving scores were posted individually on a vertical scale along with their names. Each week the deliverer with the highest average performance was rewarded with a free vehicle-maintenance coupon. Turn-signal use among drivers at Store A (n = 24) increased 22 percentage points from baseline to the intervention phase which was maintained into the withdrawal phase. Complete intersection stopping among drivers at Store B (n = 21) increased 17 percentage points from baseline to the intervention phase and was maintained into the withdrawal phase. Although the winners of the weekly competitions had the greatest increase in performance, non-winners also increased their instances of safe driving during the competition.

ARTICLE

The onset of electronic commerce, paired with an increasingly busy public, will create increases in the home delivery of products by an estimated 16% annually (Niemira, 1998). Shopping on the Internet alone is a \$5.8 billion industry and this results in many more products being delivered to homes. Additionally, an estimated \$18 billion is spent on express delivery to businesses (Niemira, 1998). Similarly, the home delivery of food products continues to grow. For example, in 1997, one national pizza delivery corporation grew at a rate of 12.5% with revenues of \$1.04 billion (Cravin, 1998). Given these data, it is expected that the number of individuals involved in delivering products as a full- or part-time occupation will continue to increase. A primary task of these employees is driving on public roads. As a result, vehicle crashes represent a significant health and productivity concern for delivery-oriented organizations, their employees, and host communities.

Pizza delivery is a particularly hazardous occupation. These drivers typically are part-time male employees, aged 18-24, with limited training and experience. They represent the gender and age group with the greatest percentage of vehicle crashes (NHTSA, 1998), and they drive for their job when traffic is heavy or when visibility is limited (i.e., 5 p.m.-2 a.m.). Plus, some pizza deliverers are paid on commission per pizza sold plus minimum wage. This translates into a contingency whereby riskier driving leads to more money. In fact, pizza deliverers are four times more likely to be involved in a vehicle crash than the rest of the population (Meagher, 1989).

Pizza deliverers are prime candidates for behavioral intervention because of the extraordinary risk of this occupation. Since they are readily observable many times during a shift, it is possible to track certain driving practices of these individuals, via license plate numbers, in a time series fashion. It is also possible to post the results of these observations publicly and conduct a competition between drivers for the best driving percentages. This was the intervention focus of the current field study.

APPLICATIONS OF FEEDBACK TO IMPROVE OCCUPATIONAL SAFETY

Behavioral feedback strategies have been applied frequently in organizational settings to improve individual and group performance (see Alvero, Bucklin, & Austin, 2001; Balcazar, Hopkins, & Suarez, 1986; Geller, 1998; and Locke & Latham, 1990, for reviews). Similarly, a number of studies have demonstrated the benefits of using individual feedback to increase safety-related behavior. Vehicle crashes among police patrollers decreased significantly when individual behavioral feedback was available from in-vehicle tachograph recorders (Larson et al., 1980). Individual feedback techniques were also effective at increasing safe work

practices in a coal mine (Rhoton, 1980), among nurses (Babcock, Sulzer-Azaroff, Sanderson, & Scibak, 1992; DeVries, Burnette, & Redmon, 1991), and among the staff in a group home (Alavosius & Sulzer-Azaroff, 1986).

A recent study by Ludwig, Geller, and Clarke (2000) added individuals' names and turn-signal use percentages to ongoing group feedback. A significant increase in turn-signal use over the group feedback phase occurred when the individual feedback was added to the public graph. Likewise, a review of 21 studies by Nordstrom, Lorenzi, and Hall (1991) concluded that public posting of individualized performance feedback was a very effective intervention technique in work settings. This included applications in factories (Emmert, 1978; McCuddy & Griggs, 1984; Wikoff, Anderson, & Crowell, 1982), schools (Gross & Ekstrand, 1983) and hospitals (Welsh, Ludwig, Radiker, & Krapfl, 1973).

For maximum impact, feedback should be salient and imply a relevant consequence (Geller, 1996). Therefore, combining feedback with a relevant contingency could make it more influential (Balcazar, Hopkins, & Suarez, 1986). One such contingency is a competition whereby participants receive rewards for the best behavioral performance. Kortick and O'Brien (1996) demonstrated that a competition between teams of package delivery employees (i.e., "The world series of quality control") could be paired with feedback and tangible rewards to improve the quality of shipping performance. More specific to behavioral safety, Haynes, Pine, and Fitch (1982) showed that the combination of feedback, team competition, and rewards decreased the rate, severity and costs of accidents among urban transit operators. The present study investigated the impact of a competition between *individuals* in a safe-driving contest that publicly posted individual feedback and provided weekly rewards to the participant who had the best behavioral feedback.

METHOD

Participants and Settings

Pizza deliverers at five nationally franchised restaurants (two intervention and three control sites) were observed departing for and arriving from their deliveries. Informed consent was obtained from participants at the two intervention sites. All vehicles used for deliveries were owned by the employees. At the time of this study, North Carolina and Virginia had laws requiring complete stops at all intersections entering public roads, as well as a law requiring turn-signal use for all turns on public roads. The law did not explicitly require turn-signal use while turning from parking lots onto public roads. In this situation, turn-signal use is considered a courtesy in the statute. Parking lots from all five stores were connected to roads posting a 35mph speed limit.

Three stores were located in a town in the southeastern U.S. with a population of 35,000 and were franchises of three different national pizza companies that offer delivery. Stores A and B were exposed to the

intervention, while Store C served as an in-town control. Store A's pizza deliverers exited onto a four-lane road from a parking lot shared with one other business. Store B's pizza deliverers exited onto a two-lane road from a parking lot shared with two other businesses. Most deliverers leaving Store B proceeded through an intersection with a four-way stop. Store C's pizza deliverers exited onto a four-lane road from a parking lot used solely by the restaurant. All three stores were within one mile of a state university serving 13,000 students.

Two additional stores were located in another town in the southeastern U.S. with a population of approximately 40,000. Stores D and E were franchised from the same national chains as Stores A and B, respectively, and served as same-company controls. This was done to control for any events (e.g., policy statements, training courses) within the company that could account for changes in safety practices among deliverers. Store D's pizza deliverers exited onto a two-lane road and shared a parking lot with one other business. Store E's pizza deliverers had their own parking lot and chose from two exits that led to a four-lane road. Both stores were within one mile of a state university serving 25,000 students.

Behavioral observations of civilians (i.e., non-pizza deliverers) were also conducted at all five observation sites. Therefore, this study had non-intervention control samples (a) within the same town, (b) within the same company, (c) with similar road conditions (i.e., Store E similar to Store A, and Store D similar to Store B), and (d) of a civilian population.

Drivers at Stores A and D earned minimum wage (i.e., \$5.50 per hour at the time of the study) and received 29¢ a mile. Drivers at Store B and E earned minimum wage and received 6% of the sales they delivered. Drivers at Store C earned minimum wage plus \$1 for every order they delivered. All deliverers earned extra money through customer tips.

Observation Procedure

Data were collected by trained observers using a behavioral checklist format applied extensively for this type of field research (Ludwig & Geller, 1991, 1997, 2000). Behavioral observations took place during peak business hours (i.e., lunch time, 11:30-1:30 p.m.; dinner time, 5:00-8:00 p.m.; and evening, 9:00-11:00 p.m.) from hidden positions overlooking the parking area of each pizza store. Data collectors were trained extensively at conducting field observations, and were blind to the scheduling and assignment of the intervention conditions. To assess interobserver reliability, observations were recorded independently by two research assistants for approximately 30% of the observation sessions.

Deliverers' use of turn signals and safety belts were recorded using a simple "yes" or "no" coding. Intersection stopping was coded as a "yes" when the vehicle's wheels stopped moving. Additionally, observers recorded vehicle license plate numbers and noted whether the driver was departing for or returning from a delivery.

Experimental Design

This study used an AA'BA multiple baseline design across two stores with non-intervention control stores and civilians. Prior to the intervention, baseline observations of pizza deliverers' complete intersection stops, turn-signal use, and safety-belt use were conducted for 7 consecutive weeks. During the last week of baseline (A') obtrusive in-store observations were conducted. A competition-based intervention was then conducted for 5 weeks followed by a 2-week withdrawal phase.

After an analysis of the baseline data, it was determined that the lowest and most stable behavior at Store A was turn-signal use, and the lowest and most stable behavior at Store B was intersection stopping. Therefore, turn-signal use was selected as the target behavior at Store A, and complete stopping became the target behavior at Store B.

In-Store Observations. One week before and during the entire intervention phase, data were collected obtrusively inside Stores A and B. This helped to preserve the anonymity of the observers collecting ongoing research data from hidden external locations. Trained observers were stationed at the front window overlooking the store's parking lot and used the same behavioral checksheet as the external hidden data collectors. In-store observers were not allowed to show employees their checksheets but if asked they were trained to respond "I am monitoring this store's [targeted behavior]." Observers stationed inside were never scheduled to collect data from the external locations. The in-store observations ceased at the end of the intervention phase, while the unobtrusive behavior observations outside the store continued throughout the three phases.

Intervention. The beginning and end of the intervention phase were staggered across the two experimental stores. The intervention was initiated first at Store A, and then one week later at Store B. At the beginning of the intervention the following information was posted in the store and put in each employee's mailbox:

- Weekly averages of [TARGETED BEHAVIOR] for each deliverer will be posted each Friday.
- The deliverer with the highest percentage of [TARGETED BEHAVIOR] each week will be awarded a prize.
- Several prizes (automotive-related items and services, such as free oil changes, valued between \$15 and \$20) are available and each week the winner will be awarded a prize.
- In addition to weekly winners, there will also be a GRAND PRIZE WINNER at the end of the contest. The GRAND PRIZE will be worth approximately twice as much as the weekly prizes (automotive related items & services, such as wheel alignment, valued between \$40 and \$50).
- Prizes will be distributed by your store manager. Deliverers can win more than one prize. Ties will be decided by random drawing.

A feedback poster was placed in the store to display deliverers' weekly occurrence percentages for the target behavior (i.e., turn-signal use at Store A and complete stops at Store B). Individual drivers' names were posted next to each percentage point. All the deliverers' names were listed vertically, corresponding to the scale for the targeted behavior on the y-axis of the graph (see Figure 1).

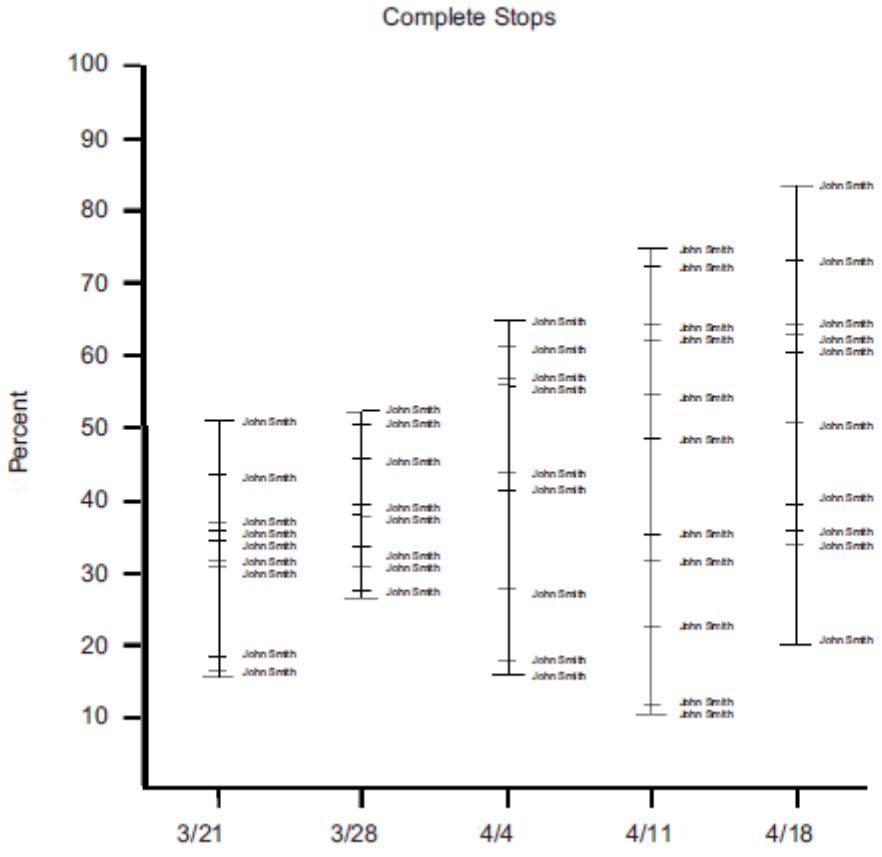
The individual pizza deliverer who had the highest percentage for the week, as shown by the highest name and data point on the graph, received a reward certificate worth approximately \$25. Individual feedback was posted once a week for four weeks, and one person received a reward certificate per week at each store.

Posted next to the weekly individual feedback graph was a cumulative graph displaying each pizza deliverer's overall performance of the target behavior over the course of the 4-week intervention. This graph added each weekly average for the individual pizza deliverers to their summed averages of the prior weeks. The deliverer at each store who had the highest cumulative average by the end of the 4-week intervention received a grand prize valued approximately \$50. See Figure 2 for an example the cumulative competition signs.

At the end of the 4-week intervention all intervention graphs were removed from Stores A and B and the in-store data collection ended. Data collection continued outside the store for a 4-week withdrawal phase.

Rewards. Two months prior to the start of baseline observations, pizza deliverers ($n = 14$) at Stores A and B were asked what they would consider an appropriate reward, other than money for performing their duties well. During this reinforcer survey (Daniels, 1989) most deliverers stated that their personal vehicle experienced a significant amount of wear during the execution of their deliveries. Therefore, they felt the most appropriate benefit the organization could extend was help in maintaining and repairing their vehicles. Therefore, rewards were solicited from local merchants specializ-

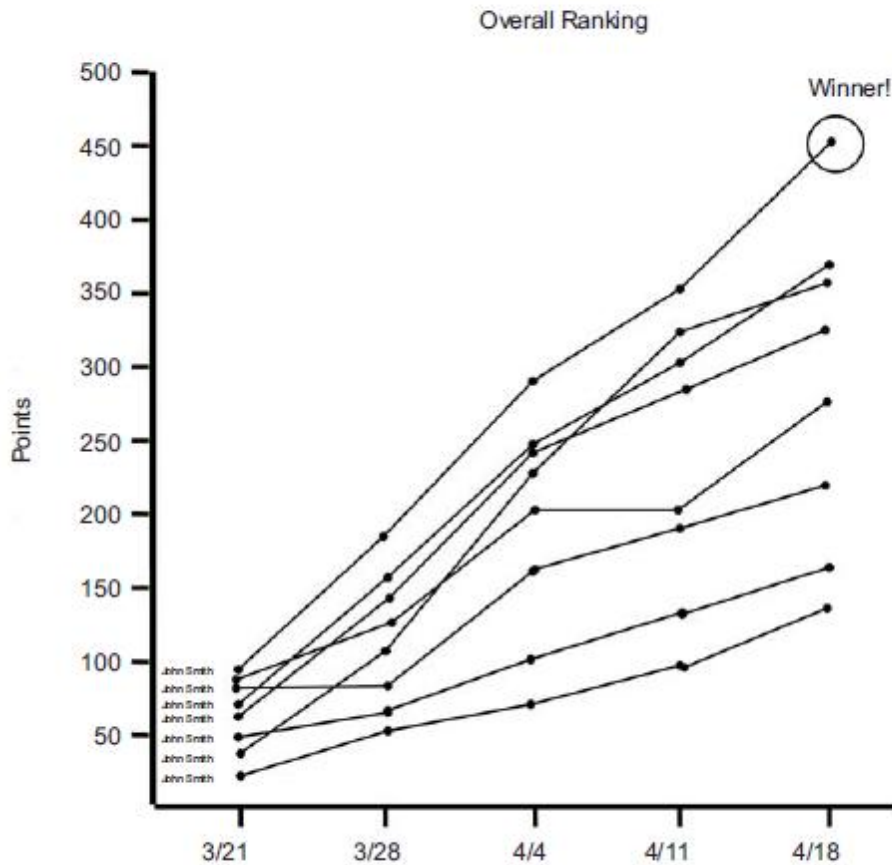
FIGURE 1. Sample weekly feedback graph displaying individual means.



ing in car care. These were in the form of coupons a pizza deliverer could redeem at the merchant store. Examples of rewards included free oil changes, car washes, wiper blades, and tire rotations. The deliverer receiving the reward could choose from any of the goods and services available. The tire rotation was the most costly of the rewards and was only available as the grand prize.

In return for their donation of goods and/or services valued around \$25, the merchants were recognized on a flyer pasted on top of outgoing pizza boxes after the data collection for this study had ceased. The flyer stated that, along with the pizza deliverers at the restaurant, the listed merchants wanted the community to drive safely. The flyers were attached to approximately 3,000 boxtops of delivered pizzas.

FIGURE 2. Sample cumulative feedback graph.



RESULTS

Over 103 days and across five observation sites, a total of 17,369 behavioral observations (including turn-signal use, safety-belt use, and complete intersection stops) were collected during 341 observation sessions. An average of 51 behavioral observations were recorded per session at each site. A total of 82 different pizza deliverers were observed at least six times per phase (Store A, $n = 24$ deliverers, 4,202 behaviors observed; Store B, $n = 21$ deliverers, 4,387 behaviors observed; Store C, $n = 10$ deliverers, 1,545 behaviors observed; Store D, $n = 14$ deliverers, 1,983 behaviors observed; Store E, $n = 13$ deliverers, 1,493 behaviors observed).

Data from pizza deliverers who were not observed at least six times during each of the three phases were deleted from the data analysis. This procedure helped assure that only deliverers who were tracked throughout the whole study contributed to the database and that there were sufficient observations to arrive at a valid estimate of performance. This criterion resulted in the elimination of the data from 86 deliverers.

Also, a total of 3,757 observations of civilian driving behaviors were conducted across all observation sites.

In-store observers were present during only 11.5% of the out-of-store observation sessions. Data collected during this overlap were not included in the primary data analysis. Instead, unless otherwise noted, the data presented below came only from data collected by the hidden observers outside the store when there were no in-store observers present. Behavioral means from the in-store observations are presented alongside the out-of-store data in Tables 1-3.

Interobserver Reliability

Interobserver reliability was assessed during 99 observation sessions at Stores A and B, balanced across the three phases. This represented

TABLE 1. Mean percentage of turn-signal use across baseline, intervention, and withdrawal phases for all participating stores and civilian observations. All percentages were derived from external observations except where in-store observations appear in italic.

Site (Target Behavior)	Baseline	Intervention	Withdrawal
Store A (Turn Signal)	35.8	57.6	52.8
<i>Store A In-Store</i>		76.5	
Store B (Compl. Stop)	25.1	28.7	11.9
<i>Store B In-Store</i>		34.1	
Store C (In-Town Control)	43.4	44.7	41.1
Store D (Corporate Ctrl.)	28.9	34.6	25.3
Store E (Corporate Ctrl.)	40.6	33.9	40.2
Civilians	38.1	34.5	30.3

TABLE 2. Mean percentage of complete intersection stops across baseline, intervention, and withdrawal phases for all participating stores and civilian observations. All percentages were derived from external observations except where in-store observations appear in italic.

Site (Target Behavior)	Baseline	Intervention	Withdrawal
Store A (Turn Signal)	35.8	29.0	48.7
<i>Store A In-Store</i>		41.5	
Store B (Compl. Stop)	14.1	30.6	20.9
<i>Store B In-Store</i>		33.6	
Store C (In-Town Control)	46.5	46.3	47.4
Store D (Corporate Ctrl.)	14.3	13.7	14.6
Store E (Corporate Ctrl.)	17.0	9.0	13.1
Civilians	38.0	31.4	32.1

TABLE 3. Mean percentage of safety-belt use across baseline, intervention, and withdrawal phases for all participating stores and civilian observations. All percentages were derived from external observations except where in-store observations appear in italic.

Site (Target Behavior)	Baseline	Intervention	Withdrawal
Store A (Turn Signal)	59.3	72.7	85.5
<i>Store A In-Store</i>		70.6	
Store B (Compl. Stop)	63.7	57.1	52.6
<i>Store B In-Store</i>		63.1	
Store C (In-Town Control)	74.1	70.2	57.1
Store D (Corporate Ctrl.)	47.0	35.9	41.9
Store E (Corporate Ctrl.)	38.8	44.7	52.2
Civilians	63.0	52.8	63.5

29% of all sessions. Interobserver agreement percentages were calculated by dividing the total number of observations agreed upon by two independent observers for a particular data category (i.e., complete

stops, safety-belt use, and turn-signal use) by the total number of agreements and disagreements, and multiplying the result by 100%. The percentages for days when reliability data were collected were then averaged to give overall interobserver reliability estimates. Overall reliability (or percentage agreement) was 89% for observations of complete stops (range: 79%-100%), 96% for safety-belt use (range: 88%-100%), and 88% for turn-signal use (range: 81%-100%). There were no notable differences in reliability results across the five experimental sites.

Behavioral Covariation

At Stores A and B, baseline correlations of deliverers' safety-belt use, turn-signal use, and complete intersection stops were conducted. These correlations during the baseline phase were .53 ($n = 43$, $p < .05$) between safety-belt and turn-signal use, .44 ($n = 43$, $p < .05$) between complete intersection stops and turn-signal use, and .10 ($n = 43$, $p > .10$) between safety-belt use and complete intersection stops. Correlations between behaviors decreased during the intervention phase to .09 ($n = 43$, $p > .10$) between safety-belt and turn-signal use, .13 ($n = 43$, $p > .10$) between complete intersection stops and turn-signal use, and .219 ($n = 43$, $p > .10$) between safety-belt use and complete intersection stops. Correlations remained low during the withdrawal phase with a .18 ($n = 43$, $p > .10$) between safety belt and turn-signal use, .13 ($n = 43$, $p > .10$) between complete intersection stops and turn-signal use, and .08 ($n = 43$, $p > .10$) between safety-belt use and complete intersection stops.

Turn-Signal Use

There were 6,535 observations of turn-signal use, averaging 19.2 observations per session (range: 8-52 observations) at each site (Store A = 1,546 observations, Store B = 1,612 observations, Store C = 579 observations, Store D = 782 observations, Store E = 595 observations, Civilians = 1,421 observations). Figure 3 depicts daily mean turn-signal use at Stores A and B within baseline, intervention, and withdrawal phases. Table 1 lists mean turn-signal use per phase for each site (including the two experimental stores, the three control stores, and civilian observations). Additionally, mean turn-signal use recorded by in-store observers during the intervention at the two experimental stores appear in Table 1.

At Store A, where the intervention targeted turn-signal use, the target behavior increased from a baseline mean of 36% to an intervention mean of 58%. During the withdrawal phase, mean turn-signal use at Store A decreased slightly to a mean of 53% but remained substantially above baseline. Turn-signal use at Store B did not increase from baseline when its deliverers received the complete stopping intervention. However, there was a downward trend in turn-signal use during the last two weeks of the intervention that continued into the withdrawal phase from a mean of 29% during the intervention phase, to 12% during withdrawal. Turn-signal use did not vary as a function of the experimental phases at the same town

control (i.e., Store C), at either of the same-corporation controls (i.e., Stores D and E), or in the civilian sample. Also, an examination of Figure 3 suggests that turn-signal use at Store A did not increase above baseline levels with the onset of the in-store observations (depicted by an arrow) during the last week of the baseline phase.

Complete Intersection Stops

There were 5,625 observations of complete intersection stopping, averaging 16.5 observations per session (range: 7-44 observations) for each site (Store A = 1,445 observations, Store B = 1,459 observations, Store C = 551 observations, Store D = 625 observations, Store E = 441 observations, Civilians = 1,104 observations). Figure 4 depicts daily complete intersection stopping at Stores A and B across baseline, intervention, and withdrawal phases. Additionally, Table 2 lists complete intersection stopping means across all phases for each store (including the two experimental stores, the three control stores, civilian observations, and in-store observations at Stores A and B). At Store B, where the intervention targeted complete intersection stopping, the target behavior increased from a mean of 14% during baseline to an intervention mean of 31%. Then during withdrawal, mean complete intersection stopping at Store B decreased to 21%.

The complete intersection stopping mean in Table 2 at Store A suggests no spread of effect during the intervention phase, although complete stopping did increase prominently during the withdrawal phase. However, further inspection of Figure 4 reveals a trend upward during the last two weeks of the intervention continuing into the withdrawal phase. Intersection stopping was quite stable across the phases at the same town control (i.e., Store C), the same-corporation controls (i.e., Stores D and E), and among the civilian sample. Also, an examination of Figure 4 suggests that complete intersection stopping did not increase at Store B above baseline

FIGURE 3. Percent turn-signal use per day across baseline, intervention, and withdrawal phases at Stores A and B. Dashed vertical lines represent phase changes and non-dashed horizontal lines within phases represent phase means. An arrow depicts the onset of in-store observations.

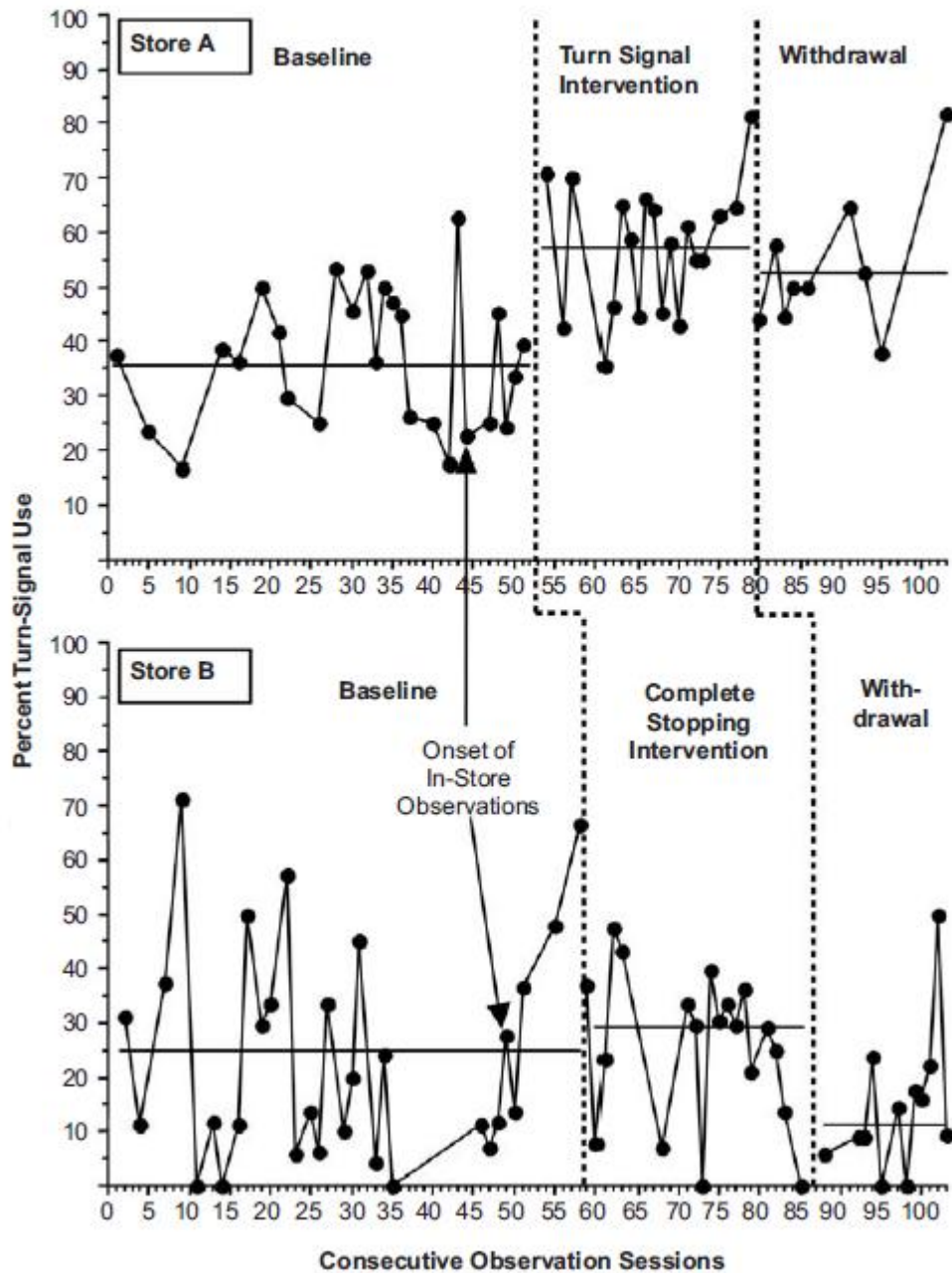
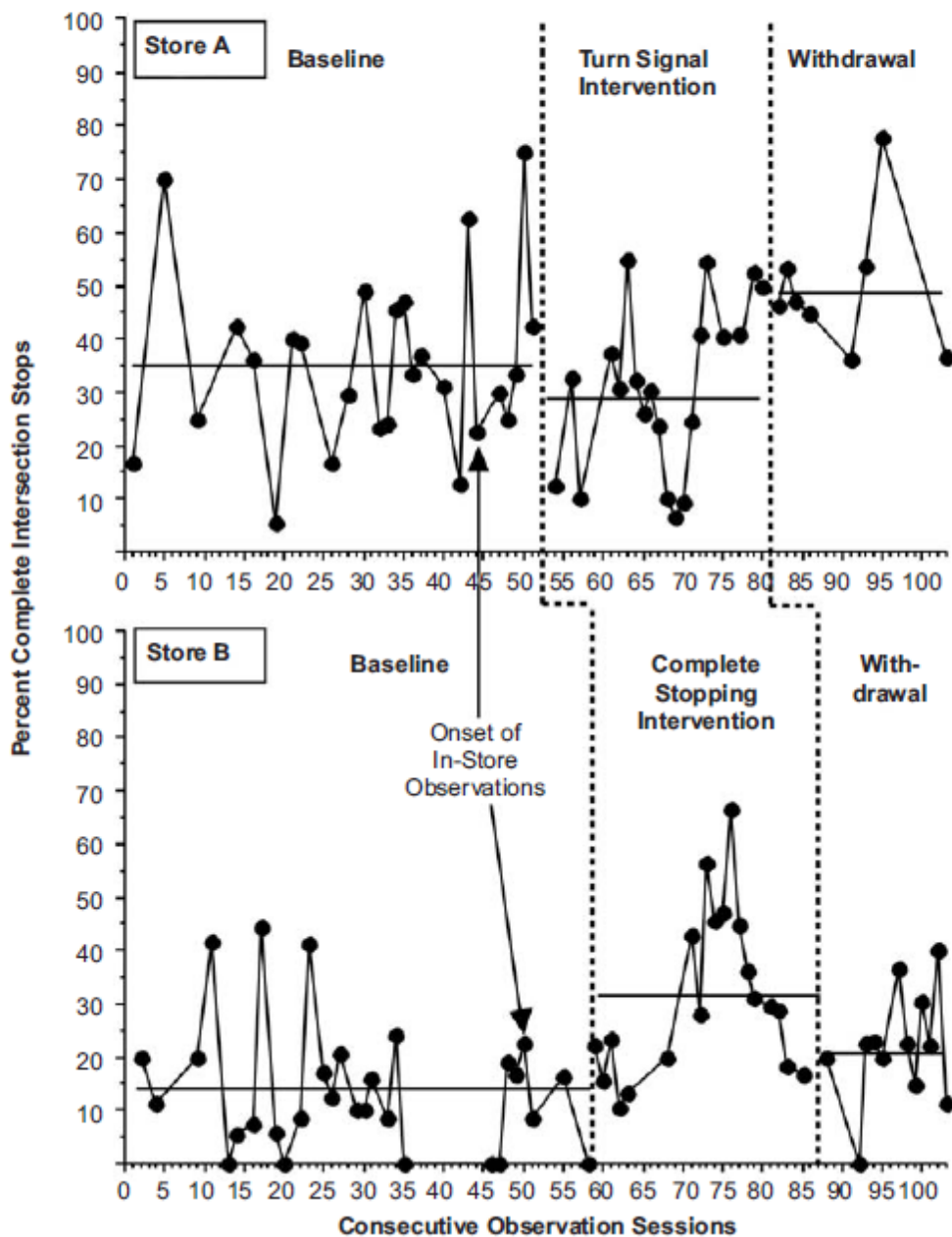


FIGURE 4. Percent complete intersection stops per day across baseline, intervention, and withdrawal phases at Stores A and B. Dashed vertical lines represent phase changes and non-dashed horizontal lines within phases represent phase means. An arrow depicts the onset of in-store observations.



levels with the onset of the in-store observations (depicted by an arrow) during the last week of the baseline phase.

Safety-Belt Use

There were 5,209 observations of safety-belt use averaging 15.3 observations per session for each site (Store A = 1,213 observations, Store B = 1,316 observations, Store C = 415 observations, Store D = 576 observations, Store E = 457 observations, Civilians = 1,232 observations). Table 3 depicts safety belt use at Stores A and B as well as the three control stores and civilian observations across baseline, intervention, and withdrawal phases. At Store A, whose intervention targeted turn-signal use, safety-belt use increased from a baseline mean of 59% to an intervention mean of 73%. And, at this store mean safety-belt use during withdrawal increased further to 86%. No other notable changes in safety-belt use occurred at the other stores or among the civilian sample.

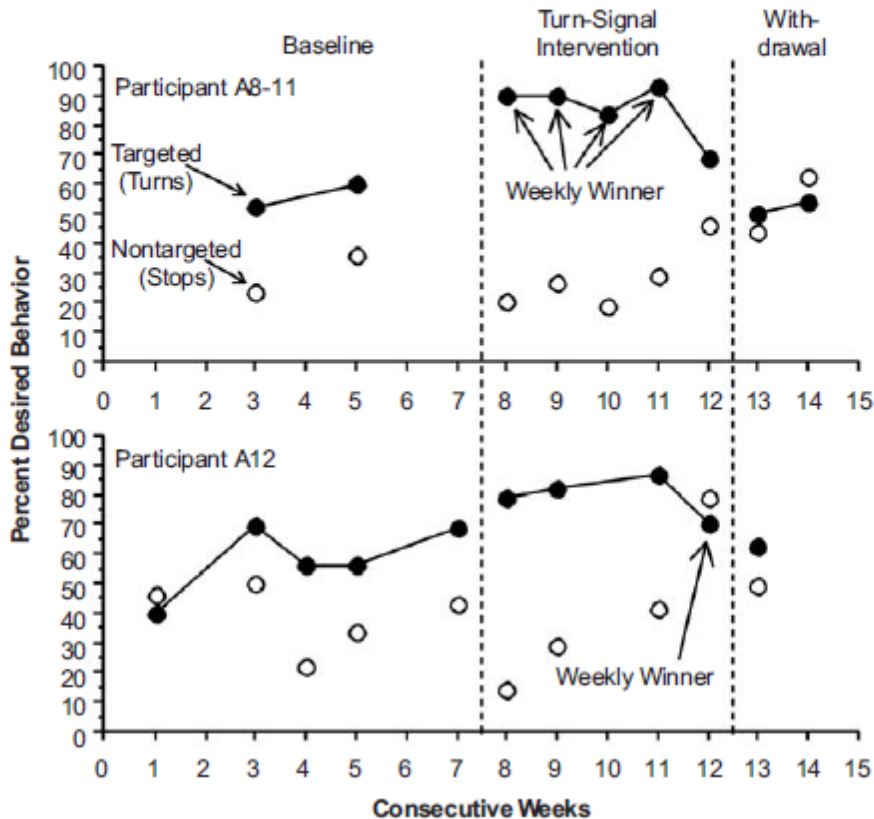
Weekly Winners

Figure 5 shows weekly behavioral means for the individuals who won a prize for having the highest turn-signal means at Store A for a given week. The individual had to be observed at least 6 times in a week to display a data point for that week. Participant A8-11 (participants were coded by store and the weeks they won) used his turn signal (depicted by filled circles) 52% and 60% of the time during baseline. He then won the first four weeks of the contest with weekly means of 90%, 89.5%, 84%, and 93%. The last week was won by Participant A12 who had a turn-signal mean of 70%. Participant A12 and Participant A8-11's turn-signal means were close over the course of the entire contest and both decreased their turn-signal use during the final week of the contest and into the withdrawal phase.

It is noteworthy that Participant A8-11's percentage of complete stops (depicted by open circles in Figure 5) trended upward toward the end of the turn-signal intervention. Likewise, Participant A12's percentage of complete stops also showed a distinctive upward trend during the turn-signal intervention.

Figure 6 shows weekly behavioral means for individuals who won a prize for having the highest complete stop means at Store B for a given week. The individual who won the first week of the contest (week 9) is not shown because she was never observed during baseline and subsequently was not used in any data analysis. Participant B10 did not show a notable

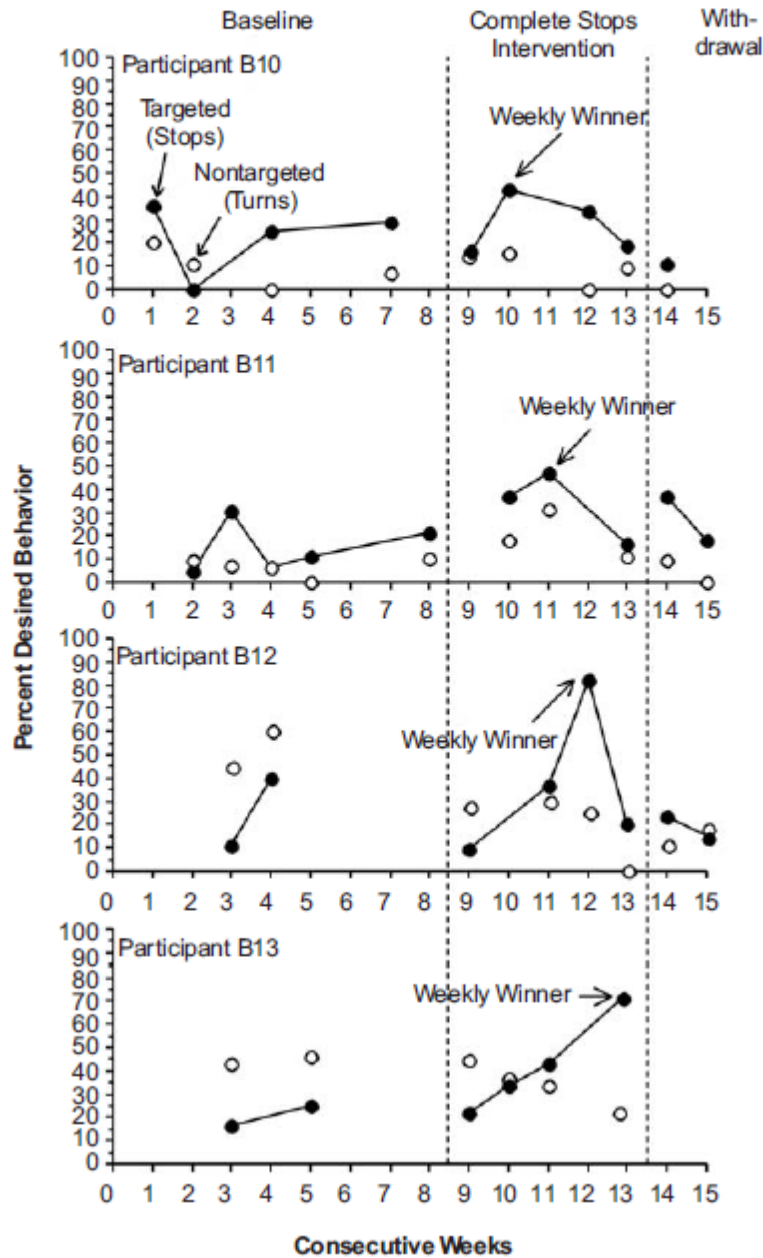
FIGURE 5. Percent desired behaviors per week for the two reward winners at Store A. Filled circles represent the targeted behavior (turn-signal use) and open circles represent a non-targeted behavior (complete stops). Participants are coded by store (i.e., "A") and the weeks in which they won the reward. Additionally, the weeks the rewards were won are noted with an arrow.



overall increase in his complete stops (depicted by filled circles) during the intervention but did increase from a mean of 14% during the first week of the contest (Week 9) to a winning second week mean of 43%.

During weeks 10 and 11, Participant B11, increased his percentage of complete stops well over baseline. He won on Week 11 with a mean of 43% complete stops. Participant B12 started the contest with a complete stop mean of 9% during Week 9 and continued to improve until he won with a mean of 84% in Week 12. Participants B10, B11, and B12 all showed declines in the weeks following their wins. Finally, Participant B13 showed a steady increase from a complete stop mean of 22% during

FIGURE 6. Percent complete intersection stops per week for the four reward winners at Store B. Filled circles represent the targeted behavior (complete stops) and open circles represent a non-targeted behavior (turn-signal use). Participants are coded by Store (i.e., "B") and the weeks in which they won the reward. Additionally, the weeks the rewards were won are noted with an arrow.



Week 9, consistent with his baseline responding, to a Week 13 mean of 71% complete stops. No weekly withdrawal observations were available for Participant B13.

Participants B10 and B11 showed conspicuous decreases in their complete stops after winning a week of the competition. To investigate this further, Figure 7 and Figure 8 was prepared to show cumulative responding of complete stops by Participant B10 and Participant B11, respectively.

Participant B10 came to a complete stop only 3 out of the next 32 opportunities we observed after his winning week. Participant B11 did not come to a complete stop for the next 23 opportunities we observed after his winning week.

Weekly Winners vs. Non-Winners

The differences in targeted behaviors between the winners and the non-winners are presented in Figure 9. At Store A the two winners showed a dramatic increase in turn-signal use over baseline. The 22 non-winners at Store A also showed an increase over baseline as a result of the intervention, albeit not as dramatic. At Store B, both the four winners (excluding Participant B9) and sixteen non-winners showed an increase during the first three weeks of the intervention (i.e., Weeks 9-11). After this point, winners and non-winners gradually decreased back to baseline levels. There were no differences in non-targeted behaviors between the winners and non-winners over the course of the study.

DISCUSSION

At Store A, an immediate increase in turn-signal use occurred at the onset of the intervention. This increase in the target behavior was sustained over the course of the intervention and maintained after the intervention was withdrawn. In contrast, complete intersection stops (Store B's intervention target) did not increase immediately at the beginning of Store B's intervention. Instead, a prominent increase in complete intersection stops was not realized until after the second week of the intervention where it spiked at 50 percentage points above baseline levels and then returned back to baseline near the end of the intervention phase where it stayed during the withdrawal phase.

Winners

FIGURE 7. Cumulative graph of Participant B10 complete intersection stopping. The week when he won his reward is noted with an arrow.

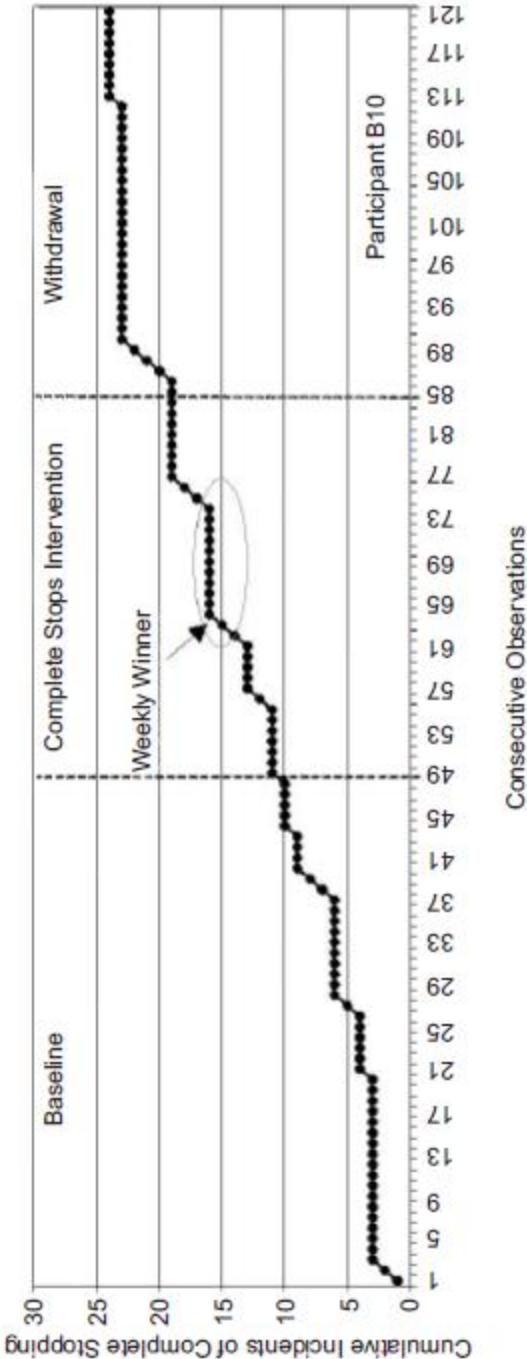


FIGURE 8. Cumulative graph of Participant B11 complete intersection stopping. The week when he won his reward is noted with an arrow.

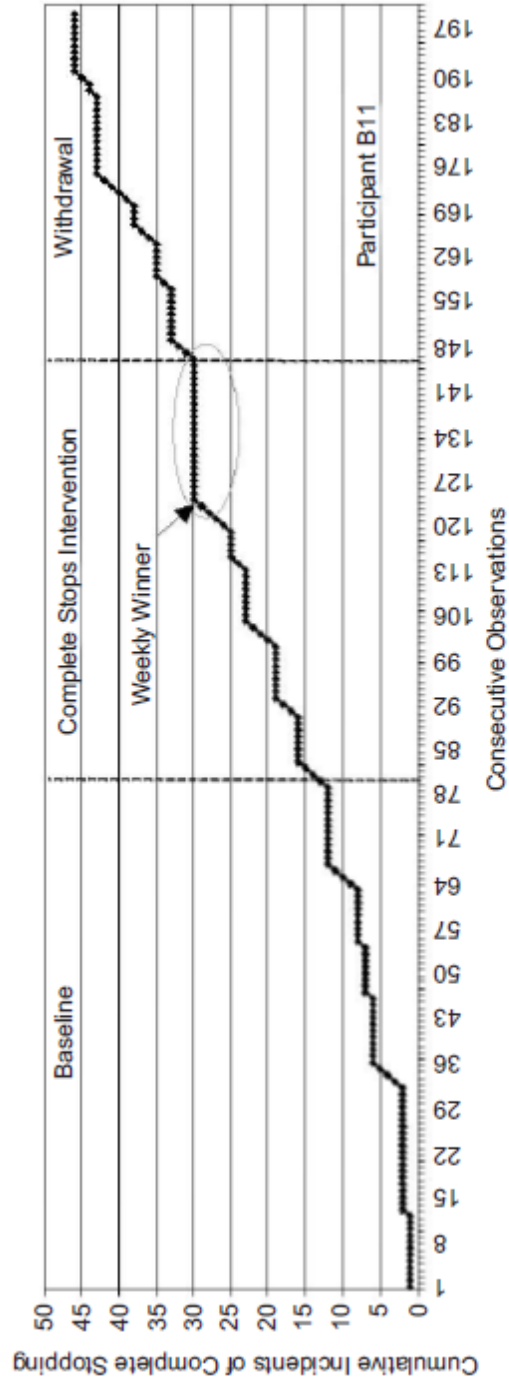
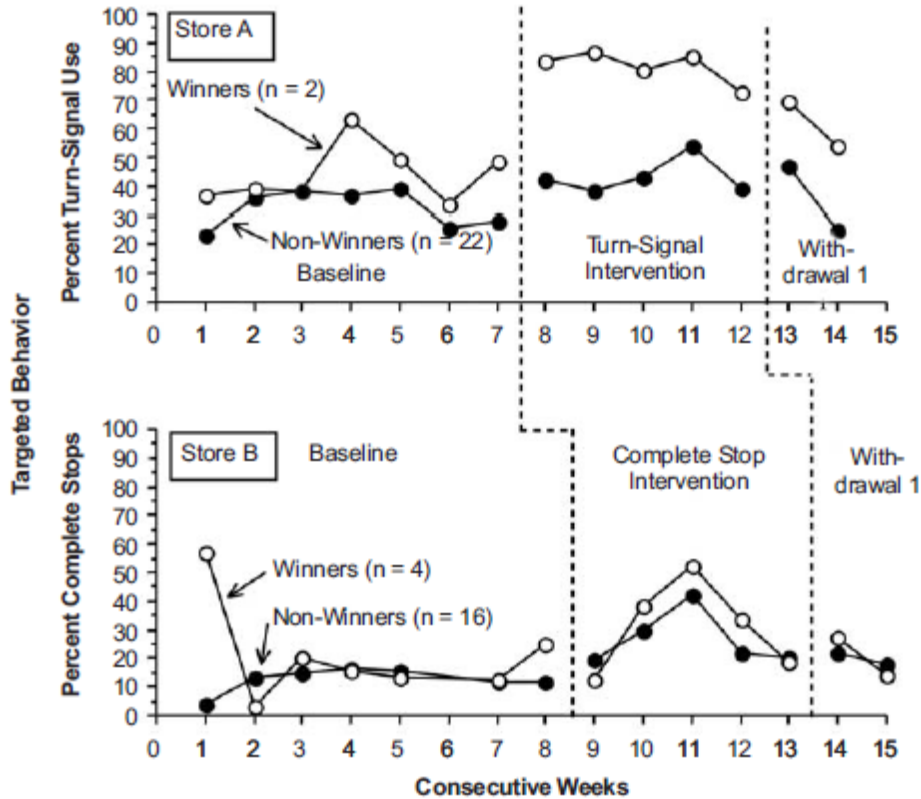


FIGURE 9. Percent turn-signal use (Store A) and complete intersection stops (Store B) per week for reward winners and non-winners (i.e., targeted behaviors). Open circles represent winners and filled circles represent non-winners.



The two winners at Store A and the 4 winners at Store B (for whom data was presented) increased their target behavior responding substantially. Participants A8-11 and A12, who were competition winners, used their turn signals nearly twice as much as the rest of the deliverers at Store A during the intervention. Although the winners at Store B (i.e., Participants B10, B11, B12, and B13) did, on average come to a complete stop more than their peers, the difference was not as prominent.

It is noteworthy that many different people won rewards at Store B. And, those winners decreased their respective behavioral percentages back to baseline levels after winning. In contrast, one person won all but one of the weeks at Store A and was challenged the entire time by only one other person (who eventually won one week). In every case among the winners, the data point following winning the reward was substantially lower.

An examination of the cumulative responding of Participant B11 revealed that he did not come to a complete stop during the next 23 instances

directly after winning a week of the competition. A similar pattern was observed in Participant B10. Such a finding is suggestive of a post-reinforcement pause (Catania, 1979), although data demonstrating post-reinforcement pauses typically are associated with fixed ratio schedules. The competition contingency in the present case approximates more of a variable ratio schedule. Nevertheless, the winners at Store B, especially Participants B10 and B11, nearly stopped their incidents of complete stopping at the observed intersection almost immediately after winning the reward coupon. Such a finding warrants additional investigation, including interviews with the winners to evaluate causes of such a pause in responding.

Non-Winners

Regarding the social validity of the incentive intervention, it is important to show that the individuals who did not win were not adversely affected by the intervention. A competition-based intervention sets up a contingency not unlike a variable ratio schedule of reinforcement because the number of correct responses required for the reward is dependent on the performance of one's competitors. However, because there can only be one winner, many of the competitors did not have an opportunity to sample the contingency (Kazdin, 1994) and, therefore, most participants' responding were followed by non-reinforcement.

Nevertheless, those who never experienced a reward for winning a week of the competition still showed a substantial improvement in the targeted behaviors. This is possibly because even non-winners were exposed to the posted individual feedback. The feedback alone may have produced behavior change in this study because it (a) specified desired behavior (on the graph), (b) prompted that behavior (the graph was visible when leaving the store), and (c) may have served as an establishing stimulus (Michael, 1982) for the competition contingency (assuming the participants had prior experience with competitions). It is also possible that the intervention may have influenced social comparison and managerial actions linked to the posted individual feedback. However, these variables were not measured in the present study.

Cooperative versus Competitive Contingencies

The results of this study are consistent with research by Kortick and O'Brien (1996) who demonstrated the positive effects of competition on quality and argued that this type of competition is similar to recreational sports whereby individuals enjoyed competing against each other.

However, there are arguments against using competition as an intervention strategy. The majority of participants in a competition are non-winners and do not get an opportunity to sample the contingency. Instead, intervention designers must rely on the prompting nature of feedback charts, establishing stimuli of individuals' past experience with competitions, and/or possible vicarious reinforcement to influence non-winner responding. In addition, Daniels (1989) argues that competition

may also lead to undesirable behaviors such as backstabbing, sabotage, and withholding information (see also Tjosvold & Deemer, 1980). Finally, in the present study, the responding of many winners decreased dramatically after receiving their reward (e.g., Participants B11 and B12).

Instead, interventions that facilitate cooperation between participants may lead to better performance than competition-based interventions. In this scenario a reward is contingent on group performance surpassing a pre-determined goal (e.g., Ludwig & Geller, 1997; Ludwig, Geller, and Clarke, 2000). By way of comparison, Allison, Silverstein, and Galante (1992) demonstrated that a cooperative condition, whereby participants received money based on group performance, was more effective when paired with feedback than a competitive condition, in which only the participants with the best performance received money. Likewise, Schmitt (1987) showed that overall response rates were lower when subjects were assigned a competitive contingency rather than a cooperative contingency.

At first glance a cooperative-based intervention may seem like a fixed ratio schedule influencing performance. However, for the individual participant, the reward offered by the cooperative-based intervention is not contingent on their behavior. Participants do not have to increase responding in order to receive the reward, they only have to rely on the performance of others. Ludwig, Geller, and Clarke (2000) demonstrated that over half of their participants did not respond to a group-based goal when group feedback was presented. Such a finding has been described as “social loafing” in other areas of psychology (Latane, 1981).

Therefore, future research should be designed to compare more directly the effects of feedback paired with a competitive versus cooperative contingency with close attention to the collateral effects of social reinforcement and possible response generalization. Additionally, we suggest that individualized response patterns should continue to be examined to more closely understand the effects of rewards in the context of these different contingencies.

Spread of Effect to Non-Targeted Behaviors

Past studies have shown that some safe driving behaviors covary rather consistently (reviewed in Ludwig & Geller, 2000). During baseline a correlation between turn-signal and safety-belt use (previously documented in Ludwig & Geller, 1991), as well as between turn-signal use and complete stops was evident at the experimental stores. In contrast, no correlation existed between safety-belt use and complete stopping. These correlations evident during baseline were much weaker when measured during the intervention and withdrawal phases suggesting that this relationship was moderated by the intervention.

Response generalization occurs when non-targeted behaviors increase as a result of intervening on a targeted behavior (Ludwig & Geller, 2000;

Russo, Cataldo, & Cushing, 1981). Evidence of response generalization (Ludwig & Geller, 2000) was mixed in the present study. An increase in safety-belt use at Store A during and after the turn-signal intervention was evident. Additionally, an increase in complete stopping occurred toward the end of the turn-signal intervention at Store A concurrent with an upward trend in the targeted behavior. In contrast, safety-belt use did not increase concurrent to the complete stopping intervention at Store B.

Turn-signal use actually decreased toward the end of Store B's turn-signal intervention concurrent with a downward trend in the targeted behavior. It is noteworthy that these general trends at Stores A and B were observed in both winners and non-winners.

Participants A8-11 and A12 both increased their complete stopping notably toward the end of Store A's turn-signal intervention. Although the increase seems to happen on Participant A12's winning week there is no such relationship in Participant A8-11's non-targeted responding and winning. Only Participant B11's non-targeted responding (turn-signal use) increased during the complete stopping intervention at Store B and this increase occurred most prominently during his winning week.

Alternatively, Participant B12 and B13's non-targeted responding trended downward during the complete stopping intervention at Store B, even as they won their week. This behavioral ecological approach (Willems, 1974) to safety research, initially suggested by Sulzer-Azaroff (1978), suggests that the effectiveness of an intervention should be assessed by its impact on not only the targeted behaviors but also functionally related behaviors not targeted by the intervention.

In-Store (Obtrusive) Observation

Obtrusive, in-store observers were used in this study to provide an ostensive source for the feedback. If the source was not evident then participants may have questioned the validity of the feedback and may have changed behavior in unpredictable ways based on the high probability of hidden observers.

Most studies reveal the source of feedback during the intervention. In these studies the data used for feedback are collected by research assistants (e.g., Alavosius & Sulzer-Azaroff, 1990; Austin et al., 1996), supervisors (e.g., Anderson, Crowell, Hantula, & Siroky, 1988; Wikoff, Anderson, & Crowell, 1983), or peers (e.g., DeVries, Burnette, & Redmon, 1991) who have been observing behavior throughout the baseline phase on into the intervention and beyond. In these cases the observers may have been visible to participants of the study yet still innocuous during baseline because no information was provided as to what the observers were doing. During the intervention these observers become an SD for the feedback contingency. In other words, observation occurring in conjunction with feedback essentially specifies what behaviors are being observed, when these behaviors are being recorded for the feedback (i.e., when the observer is present) and, therefore, when correct responding will be reinforced. One cannot, therefore, separate the impact of the feedback

intervention from the impact of observation even if the observers were previously visible to the participants during baseline.

Indeed, in the present study we cannot adequately discriminate the effects of the competition and feedback intervention from the probable impact of the in-store observer. In-store observers were introduced to the milieu one week before the onset of the intervention. No immediate effects of these observers were noted in the data. However, one week may not be enough time to assess the impact of overt observation and future studies should extend this phase of the experimental design.

One advantage of the current design is that the hidden observers could monitor deliverers' driving behaviors during the intervention when the SD of the overt observer (i.e., in-store observer) was not present. Indeed, at Store A there was a 20 percentage point difference in observed turn-signal use between the data collected by external versus internal observers (taken when the other observation post was not manned). No such difference was observed in Store B's complete stopping. (It should be noted, however, that there was no reliability collected for the in-store data.) Hence, the current design allowed us to have a true baseline where there was no overt observation stimulus. The design also provided us with an independent sample of behavior during the intervention apart from the overt observations.

Future research using feedback should employ independent samples of behavior that can evaluate participant behavior when the feedback observation SD is present versus when it is absent. Additionally, inter-observer agreement data should be collected on any overt observations to assess reliability and make it possible to make more valid comparisons between covert and overt observations during feedback interventions. Finally, the introduction of the overt observation technique should take place after a period of true baseline but with enough time left before the onset of the feedback intervention to assess the independent impact of overt observation.

With that said, there may not be a need to separate the effects of feedback from the observation stimulus. In fact, a feedback intervention is already inclusive of an observation stimulus. They cannot be separated. It would be impossible to provide feedback without collecting human, mechanical, or archival observations on behavior first. Therefore, to investigate the separate impact of observation apart from the overall feedback-based intervention would be akin to calling the observation stimulus an external variable to the design, which cannot be the case when using feedback. Certainly, the role of overt observation within studies employing feedback strategies needs to be debated further among behavioral researchers.

LIMITATIONS AND CONCLUSIONS

A couple of limitations of this study should be addressed when conducting future research in this area. There was no assessment of social validity in this study. This was especially relevant because of the use of names on the feedback graphs and the competitive nature of the intervention. Secondly, participants could have been interviewed to better understand the peer and supervisor interactions that may have influenced the data. On site observations of these milieus and other controlling variables may also be worthwhile. Such data could help us understand why we found divergent trends in responding between the intervention stores.

In conclusion, this study suggests that public individualized feedback paired with competition can be used as an effective intervention strategy to increase the occurrence of safe driving practices. Moreover, it's likely such a strategy will benefit the behaviors of both those who win the competition as well as those who do not.

REFERENCES

- Alavosius, M. P., & Sulzer-Azaroff, B. (1990). Acquisition and maintenance of health-care routines as a function of feedback density. *Journal of Applied Behavior Analysis*, 23, 151-162.
- Allison, D. B., Silverstein, J. M., & Galante, V. (1992). Relative effectiveness and cost-effectiveness of cooperative, competitive, and independent monetary incentive systems. *Journal of Organizational Behavior Management*, 13, 85-112.
- Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness of and essential characteristics of performance feedback in organizational settings (1985-1998). *Journal of Organizational Behavior Management*, 21, 3-29.
- Anderson, D. C., Crowell, C. R., Hantula, D. A., & Siroky, L. M. (1988). Task clarification and individual performance posting for improving cleaning in a student-managed university bar. *Journal of Organizational Behavior Management*, 9, 73-90.
- Austin, J., Kessler, M. L., Riccobono, J. E., & Bailey, J. S. (1996). Using feedback and reinforcement to improve the performance and safety of a roofing crew. *Journal of Organizational Behavior Management*, 16, 49-75.
- Babcock, R. A., Sulzer-Azaroff, B., Sanderson, M., & Scibak, J. (1992). Increasing nurses' use of feedback to promote infection-control practices in a head-injury treatment center. *Journal of Applied Behavior Analysis* 25, 621-627.

Balcazar, F. E., Hopkins, B. L., & Suarez, Y. (1986). A critical, objective review of performance feedback. *Journal of Organizational Behavior Management*, 7, 65-89.

Catania, A. C. (1979). *Learning*. Englewood Cliffs, NJ: Prentice-Hall.

Cravin, D. R. (1998). Domino's looks abroad to deliver more growth. *Crains' Detroit Business*, 14, 3.

Daniels, A. C. (1989). *Performance management: Improving quality productivity through positive reinforcement*. Tucker, GA: Performance Management Publications.

DeVries, J. E., Burnette, M. M., & Redmon, W. K. (1991). AIDS prevention: Improving nurses' compliance with glove wearing through performance feedback. *Journal of Applied Behavior Analysis*, 24, 705-711.

Emmert, G. D. (1978). Measuring the impact of group performance feedback versus individual performance feedback in an industrial setting. *Journal of Organizational Behavior Management*, 1, 134-141.

Geller, E. S. (1996). *The psychology of safety: How to improve behaviors and attitudes on the job*. Radnor, PA: Chilton Book Company.

Geller, E. S. (1998). *Applications of behavior analysis to prevent injuries from vehicle crashes* (Revised Edition). Behavior Monographs from the Cambridge Center for Behavioral Studies, Cambridge, MA.

Gross, A. M., & Ekstrand, M. (1983). Increasing and maintaining rates of teacher praise. *Behavior Modification*, 7, 126-135.

Haynes, R. S., Pine, R. C., & Fitch, H. G. (1982). Reducing accident rates with organizational behavior modification. *Academy of Management Journal*, 25, 407-416.

Kazdin, A. E. (1994). *Behavior modification in applied settings*. Pacific Grove, CA: Brooks/Cole Publishing.

Kortick, S. A., & O'Brien, R. M. (1996). The world series of quality control: A case study in the package delivery industry. *Journal of Organizational Behavior Management*, 16, 77-93.

Larson, L. D., Schnelle, J. F., Kirchner, Jr., R., Carr, A. F., Domash, M., & Risley, T. R. (1980). Reduction of police vehicle accidents through mechanically aided supervision. *Journal of Applied Behavior Analysis*, 13, 571-581.

Latane, B. (1981). The psychology of social impact. *American Psychologist*, 36, 343-356.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.

Ludwig, T. D., & Geller, E. S. (1991). Improving the driving practices of pizza deliverers:

Potential moderating effects of age and driving record. *Journal of Applied Behavior Analysis*, 24, 31-44.

Ludwig, T. D., & Geller, E. S. (1997). Assigned versus participatory goal setting and response generalization: Managing injury control among professional pizza deliverers. *Journal of Applied Psychology*, 82, 253-261.

Ludwig, T. D., & Geller, E. S. (1999a). Behavior change among agents of a community safety program: Pizza deliverers advocate community safety-belt use. *Journal of Organizational Behavior Management*, 19, 3-24.

Ludwig, T. D., & Geller, E. S. (1999b). Paradoxical effects of repeated turn-signal policy memos: Evidence of countercontrol behaviors. *Journal of Organizational Behavior Management*, 19, 25-34.

Ludwig, T. D., & Geller, E. S. (2000). Intervening to improve the safety of delivery drivers: A systematic behavioral approach. Monograph. *Journal of Organizational Behavior Management*, 19, 1-153.

Ludwig, T. D., Geller, E. S., & Clarke, S. W. (2000). Using publicly-displayed feedback to increase turn-signal use: Examining a spread of effect to safe stops and safety-belt use. Manuscript submitted for publication.

McCuddy, M. K., & Griggs, M. H. (1984). Goal setting and feedback in the management of a professional department: A case study. *Journal of Organizational Behavior Management*, 6, 53-64.

Meagher, M. (1989, February). Death in the delivery zone (television segment). *Inside Edition*, American Broadcasting Company.

Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior*, 37, 149-155.

Niemira, M. P. (1998). The mail-order barometer. *Chain Store Age*, 74, 40.

Nordstrom, R., Lorenzi, P., & Hall, R. V. (1990). A review of public posting of performance feedback in work settings. *Journal of Organizational Behavior Management*, 11, 101-124,

Rhoton, W. W. (1980). A procedure to improve compliance with coal mine safety regulations. *Journal of Organizational Behavior Management*, 2, 243-260.

Russo, D. C., Cataldo, M. F., & Cushing, P. J. (1981). Compliance training and behavioral covariation in the treatment of multiple behavior problems. *Journal of Applied Behavior Analysis*, 14, 209-222.

Schmitt, D. R. (1987). Interpersonal contingencies: Performance differences and cost-effectiveness. *Journal of the Experimental Analysis of Behavior*, 48, 221-234.

Sulzer-Azaroff, B. (1978). Behavioral ecology and accident prevention. *Journal of Organizational Behavior Management*, 2, 11-44.

Tjosvold, D., & Deemer, D. K. (1980). Effects of controversy within a cooperative or competitive context on organizational decision making. *Journal of Applied Psychology*, 65, 590-595.

U.S. Department of Health and Human Services & Public Health Services. (1991). *Healthy people 2000: National health promotion and disease prevention objectives*. Special Research Report, Washington, DC: DHHS Publication No. (PHS) 91-50212.

Welsch, W.V., Ludwig, C., Radiker, J. E., & Krapfl, J. E. (1973). Effects of feedback on daily completion of behavior modification projects. *Mental Retardation*, 11, 24-26.

Wikoff, M., Anderson, D. C., & Crowell, C. R. (1983). Behavior management in a factory setting: Increasing work efficiency. *Journal of Organizational Behavior Management*, 4, 97-127.

Willems, E. P. (1974). Behavioral technology and behavioral ecology. *Journal of Applied Behavior Analysis*, 7, 151-165.