

*Journal of*

---

# **INDUSTRIAL TECHNOLOGY**

---

*Volume 20, Number 2 - February 2004 to April 2004*

---

## ***A Statistical Comparison of Three Root Cause Analysis Tools***

*By Dr. Anthony Mark Doggett*

*Peer-Refereed Article*

**KEYWORD SEARCH**

*Leadership  
Management  
Quality  
Research  
Sociology*



Dr. Mark Doggett is a post-doctoral fellow and instructor at Colorado State University and is an adjunct faculty member at Aims Community College. He is currently working on grants for the National Science Foundation in medical technology and the Department of Education in career and technical education. He also teaches courses in process control, leadership, and project management. His research interests are decision-making and problem-solving strategies, technical management, theory of constraints, and operations system design.

# A Statistical Comparison of Three Root Cause Analysis Tools

By Dr. Anthony Mark Doggett

To solve a problem, one must first recognize and understand what is causing the problem. According to Wilson et al. (1993), a root cause is the most basic reason for an undesirable condition or problem. If the real cause of the problem is not identified, then one is merely addressing the symptoms and the problem will continue to exist. For this reason, identifying and eliminating root causes of problems is of utmost importance (Dew, 1991; Sproull, 2001). Root cause analysis is the process of identifying causal factors using a structured approach with techniques designed to provide a focus for identifying and resolving problems. Tools that assist groups and individuals in identifying the root causes of problems are known as root cause analysis tools.

## Purpose

Three root cause analysis tools have emerged from the literature as generic standards for identifying root causes. They are the cause-and-effect diagram (CED), the interrelationship diagram (ID), and the current reality tree (CRT). There is no shortage of information available about these tools. The literature provided detailed descriptions, recommendations, and instructions for their construction and use. The literature documented processes and structured variations for each tool. Furthermore, the literature is quite detailed in providing colorful and illustrative examples for each of the tools so practitioners can quickly learn and apply them. In summary, the literature confirmed that these three tools do, in fact, have the capacity to find root causes with varying degrees of accuracy, efficiency, and quality (Anderson & Fagerhaug, 2000; Arcaro, 1997; Brown, 1994; Brassard, 1996; Brassard & Ritter, 1994; Cox &

Spencer, 1998; Dettmer, 1997; Lepore & Cohen, 1999; Moran et al., 1990; Robson, 1993; Scheinkopf, 1999; Smith, 2000)

For example, Ishikawa (1982) advocated the CED as a tool for breaking down potential causes into more detailed categories so they can be organized and related into factors that help identify the root cause. In contrast, Mizuno (1979/1988) supported the ID as a tool to quantify the relationships between factors and thereby classify potential causal issues or drivers. Finally, Goldratt (1994) championed the CRT as a tool to find logical interdependent chains of relationships between undesirable effects leading to the identification of the core cause.

A fundamental problem for these tools is that individuals and organizations have little information to compare them to each other. The perception is that one tool is as good as another tool. While the literature was quite complete on each tool as a stand-alone application and their relationship with other problem solving methods, the literature is deficient on how these three tools directly compare to each other. In fact, there are only two studies that compared them and the comparisons were qualitative. Fredendall et al. (2002) compared the CED and the CRT using previously published examples of their separate effectiveness while Pasquarella et al. (1997) compared all three tools using a one-group post-test design with qualitative responses. There is little published research that quantitatively measures and compares the CED, ID, and CRT. This study attempted to address those deficiencies.

The purpose of this study was to compare the perceived differences between the independent variables: the

cause-and-effect diagram (CED), the interrelationship diagram (ID), and the current reality tree (CRT) with regard to causality, factor relationships, usability, and participation. The first dependent variable was the perceived ability of the tool to find root causes and the interdependencies between causes. The second dependent variable was the perceived ability of the tool to find relationships between factors or categories of factors. Factors may include causes, effects, or both. The third dependent variable was the overall perception of the tool's usability to produce outputs that were logical, productive, and readable. The fourth dependent variable was the perception of participation resulting in constructive discussion or dialogue. In addition, the secondary interests of the study were to determine the average process times required to construct each tool, the types of questions or statements raised by participants during and after the process, and the nature of the tool outputs.

**Delimitations, Assumptions, and Limitations**

The delimitations of the study were that the tools were limited to the CED, ID, and CRT while participants were limited to small groups representing an authentic application of use. The limitations of the study were grounded in the statistical requirements for the General Linear Model. The experimental results reflected the efficacy of the tools in the given context. While the researcher attempted to control obvious extraneous variables during the study, participant and organizational cultural attributes, politics, and social climate remained outside the scope and control of the analysis.

The assumptions of the study were that (a) root cause analysis techniques are useful in finding root causes, (b) the identification of a root cause will lead to a better solution than the identification of a symptom, and (c) the identification of causal interdependencies is important. In addition, expertise, aptitude, and prior knowledge about the tools, or lack thereof, were assumed to be randomly distributed within the

participant groups and did not affect the overall perceptions or results. Also, the sample problem scenarios used in the study were considered as having equal complexity.

***Methodology***

The specific design was a within-subjects single factor repeated measures with three levels. The independent variables were counterbalanced as shown in Table 1, where *T* represents the treatment, *M* represents the measure, and the group observations are indicated by *O*. The rationale for this design is that it compares the treatments to each other in a relative fashion using the judgments of the participants. In this type of comparative situation, each participant serves as his or her own control making the use of independent groups unnecessary (Girden, 1992). The advantage of this design is that it required fewer participants while reducing the variability among them, which decreased the error term and the possibility of making a Type I error. The disadvantage was that it reduced the degrees of freedom (Anderson, 2001; Gliner & Morgan, 2000; Gravetter & Wallnau, 1992; Stevens, 1999).

**Measures and Instrument**

Three facilitators were trained in the tools, processes, and procedures before the experiment. They were instructed to be available to answer questions from the participants about the tools, goals, purpose, methods, or instructions. The facilitators did not provide information about the problem scenarios. They were also trained in observational techniques and instructed to intervene in the treatment process if a group was having difficulty constructing the tool or managing their

process. The activity of the facilitators was intended to help control the potential diffusion of treatment across the groups.

To ensure consistency, each treatment packet was similarly formatted with the steps for tool construction and a graphical example based on published material. Each treatment group also received the same supplies for constructing the tools. The dependent variables were measured using a twelve-question self-report questionnaire with a five-point Likert scale and semantic differential phrases.

**Procedure**

Participants and facilitators were randomly assigned to one of three groups. The researcher provided simultaneous instructions about the experiment, problem scenarios, and materials. Five minutes were allowed for the participants to review their respective scenarios and instructions, followed by a ten minute question period. The participants were then asked to analyze and find the perceived root cause of the problem. The facilitators were available for help throughout the treatment. Upon completion of the treatment, the participants completed the self-report instrument. This process was repeated until all groups applied all three analysis tools to three randomized problems. Each subsequent treatment was administered every seven days.

**Reliability and Validity**

Content validity of the instrument was deemed adequate by group of graduate students familiar with the research. Cronbach's alpha was .82 for a pilot study of 42 participants. The dependent variables were also congruent with an exploratory principle component analysis.

**Table 1. Repeated Measures Design Model**

Group	T <sub>1</sub>	M <sub>1</sub>	T <sub>2</sub>	M <sub>2</sub>	T <sub>3</sub>	M <sub>3</sub>
1	CED	O <sub>1,1</sub>	CRT	O <sub>1,2</sub>	ID	O <sub>1,3</sub>
2	ID	O <sub>2,1</sub>	CED	O <sub>2,2</sub>	CRT	O <sub>2,3</sub>
3	CRT	O <sub>3,1</sub>	ID	O <sub>3,2</sub>	CED	O <sub>3,3</sub>

Threats to internal validity were maturation and carryover effects. Other threats included potential participant bias, statistical regression to the mean, outside influence, interruptions, and interaction between participants. An additional threat was attrition of participants due to fatigue, boredom, or time constraints.

The most significant threat to external validity was the representativeness of the selected sample because of ecological components associated with qualitative research. The setting and context were congruent with a typical educational environment. Therefore, the assumption of generalizability would be based on a judgment about the similarity between the empirical design and an ideal situation in the field (Anderson, 2001). From a pure experimental standpoint, the external validity might be considered low, but from a representative design standpoint, the validity was high (Snow, 1974).

**Participants**

The participants were first and second year undergraduate students in three intact classroom sections of a general education course on team problem solving and leadership. Each group consisted of ten to thirteen participants and was primarily white males. Females accounted for 11% of the sample and minorities comprised less than 3%. The initial sample was 107 participants. The actual sample was 72 participants due to attrition and non-responses on the instrument.

**Data Analysis**

A repeated-measures ANOVA with a two-tailed .05 alpha level was selected for the study. For significant findings, post hoc *t* tests with Bonferroni adjustments identified specific tool differences based on conditions of sphericity (Field, 2000; Stevens, 1999) and calculated effect sizes (*d*) (Cohen, 1988). As in other ANOVA designs, homogeneity of variance is important, but in repeated measures, each score is somewhat correlated with the previous measurement. This correlation is known as sphericity. While repeated-measures

**Table 2. Repeated Measures Analysis of Variance (ANOVA) for Individual Question Regarding Cause Categories**

Source/Contrast	SS	df	MS	F	p
Within-Treatments	12.84	2	6.42	7.839	.001
Error	121.16	148	.819		

**Table 3. Significant Within-Group Differences for Usability Variable**

Source/Contrast	SS	df	MS	F	p
Within-Treatments	6.80	1.881*	3.62	9.156	.000
Error	54.99	139.20*	.395		

\*Huynh-Feldt adjusted df

ANOVA is robust to violations of normality, it is not robust to violations of sphericity. For violations of sphericity, the researcher used the Huynh and Feldt (1976) corrected estimates suggested by Girden (1992). All statistical analyses were performed using the Statistical Package for Social Sciences™(SPSS) software.

**Statistical Findings**

Screening indicated the data were normally distributed and met all assumptions for parametric statistical analysis. After all responses, Cronbach’s alpha for the instrument was .93. Descriptive statistical data for the dependent variables found that the mean for the CED was either the same or higher on all dependent variables with standard deviations less than one. For the individual questions on the instrument, the mean for the CED was higher on eight questions, while the mean for the ID was higher on four.

No statistical difference was found among the three tools for causality or participation. Therefore, the null hypothesis ( $H_0$ ) was retained and there does not appear to be a difference in the perceptions of the participants concerning the ability of the tools to identify causality or affect participation.

No statistical difference was found among the three tools regarding

factor relationships. Therefore, the null hypothesis ( $H_0$ ) for factor relationships was retained. However, as shown in source Table 2, there was a significant difference in response to an individual question on this variable regarding how well the tools identify categories of causes ( $F(2, 74) = 7.839, p = .001$ ). Post hoc tests found that the CED was perceived as statistically better at identifying cause categories than either the CRT ( $t(85) = 4.54, p < .001$ ) or the ID ( $t(83) = 2.81, p = .023$ ) with medium effect sizes of 0.59 and 0.47 respectively.

Using a corrected estimate, a significant statistical difference was also found for usability ( $F(1.881, 74) = 9.156, p < .001$ ). Post hoc comparisons showed that both the CED ( $t(85) = 5.04, p < .001$ ) and ID ( $t(81) = 2.37, p = .009$ ) were perceived more usable than the CRT with medium effects sizes of 0.56 and 0.53, respectively. The overall results for usability are shown in source Table 3. Therefore, the null hypothesis ( $H_0$ ) was rejected and there is a significant difference between the CED, ID, and CRT with regard to perceived usability.

The usability variable was the perception of the tool’s ease or difficulty, productive output, readability, and assessment of integrity. This variable was measured using four

questions, to which participants perceived a significant difference on three questions. The statistical results for the individual questions about usability are shown in Table 4.

Using a corrected estimate, participants responded significantly to the question regarding perceived ease or difficulty of use ( $F(1.886, 71) = 38.395, p < .001$ ). The post hoc comparison revealed a large difference between the CRT and CED ( $t(81) = 8.95, p < .001, ES = 1.15$ ) as well as between the CRT and ID ( $t(77) = 5.99, p < .001, ES = 1.18$ ). Thus, the CRT is perceived as being much more difficult than either the CED or ID.

Participants responded significantly to the question regarding perceived differences between the tools for productive problem solving activity ( $F(2, 71) = 3.650, p = .028$ ). Post hoc tests found that the CED was perceived to be statistically better at facilitating productive problem solving activity than the CRT ( $t(81) = 3.67, p = .010$ ). The result was a small-to-medium effect size of 0.38.

Participants responded significantly to the question regarding the perceived readability of the tools ( $F(2, 71) = 3.480, p = .033$ ). Post hoc tests indicated the CED was significantly preferred over the CRT for overall readability ( $t(81) = 3.41, p = .021$ ) with a small-to-medium effect size of 0.39.

**Summary of Statistical Findings**

The statistical significance of usability was primarily due to significant differences on three of the four individual questions. The large effect size between the CRT and the other tools in response to ease or difficulty of use was the dominant factor. Thus, the CRT was deemed more difficult than the other tools. The other significant statistical finding was that the CED was perceived better at identifying cause categories than either the ID or CRT. However, this finding did not change participant’s perceptions of factor relationships. Post hoc comparisons for all significant findings are shown in Table 5.

**Other Findings**

**Process Times and Counts**

The mean times for the CED, ID, and CRT were 26 minutes, 26 minutes, and 30 minutes, respectively. The CED had the smallest standard deviation at 5.59 and the CRT had the largest standard deviation at 9.47. The standard deviation for the ID was 8.94. The researcher also counted the number of factors and arrows on each tool output. On the average, both the CED and CRT used less factors and arrows than the ID, but the CRT used a third less arrows than either the CED or ID.

**Open-Ended Participant Responses**

Comments received were generally about the groups, the process, or the tools. Comments about the groups were typically about group size or the amount of participation. Comments about the process were typically complaints or comments and were either positive or negative depending on the participant’s experience. Comments about the CED and ID tools were either favorable or ambiguous. Most of the comments about the CRT concerned its degree of difficulty.

**Researcher Observations**

The researcher sorted participant questions and facilitator observations using key words in context to discover emergent themes or categories. This sorting resulted in four categorical types: process, construction, root cause, and group dynamics. Questions raised during the development of the CED were about the cause categories and if multiple root causes were acceptable. Questions about the ID were primarily about the direction of the arrows. Questions about the CRT were varied, but generally about the tool process or construction. A common question from participants among all three tools was whether their assignment was to find a root cause or to determine a solution to the problem.

Most facilitator observations were about group dynamics or the overall process. Typical facilitator comments concerned the degree of participation or leadership roles. Facilitator observations were relatively consistent regardless of the tool used, except for the CRT. Those observations were different in that the construction observations were much higher and there were no observations concerning difficulty with root causes.

**Table 4. Repeated Measures Analysis of Variance (ANOVA) for Individual Questions Regarding Usability**

Source/Contrast	SS	df	MS	F	p
<i>Productive Problem Problem-Solving Activity</i>					
Within-Treatments	4.01	2	2.01	3.650	.028
Error	77.99	142	.549		
<i>Ease or Difficulty of Use</i>					
Within-Treatments	49.37	1.866*	26.46	38.395	.000
Error	91.30	132.47*	.689		
<i>Readability</i>					
Within-Treatments	4.70	2	2.35	3.480	.033
Error	95.96	142	.676		

\*Huynh-Feldt adjusted df

The tool outputs were also qualitatively evaluated for technical accuracy, if a root cause was identified, and the integrity of the root cause. The technical accuracy of the tool was evaluated based on the (a) directionality of the cause and effect relationships; (b) specificity of the factors; (c) format, and (d) use of conventions. Root cause was evaluated by whether the group, through some means, visual or verbal, identified a root cause during the treatment. Groups that were unable to identify a root cause either disagreed about the root cause or indicated that a root cause could not be found. The integrity of the root cause was evaluated based upon specificity and reasonableness. Specificity was defined as something that could be specifically acted upon, whereas reasonableness was defined such that someone not familiar with the problem would state that the root cause seemed rational or within the bounds of common sense.

**Summary of Other Findings**

A summary of the questions, observations, and tool output evaluations is shown in Table 6. The technical accuracy of both the CED and the ID were high, whereas the technical accuracy of the CRT was mixed. Groups using the CED were seldom able to identify a specific root cause while the groups using the ID did better. Because groups using the CED could not identify a specific root cause, the integrity of their selections also suffered. Only one CED group found a root cause that was both specific and reasonable. While the ID groups found a root cause most of the time, the integrity of their root causes was mixed. In contrast, CRT groups found a root cause most of the time with high integrity in over half the cases. In addition, the CRT groups were able to do this using less factors and arrows as the other tools.

The number of root cause questions and observations for the CED was twice as high as the ID and ten times greater than the CRT. Conversely, the CRT generated more process and construction questions/observations than the ID or CED. In summary, the CED pro-

**Table 5. Post Hoc T-Tests with Bonferroni Adjustment**

Dependent Variable Construct	CED		ID		CRT		df	t	p	d
	M	SD	M	SD	M	SD				
<b>Relationships</b>										
<b>Identified Cause Categories</b>										
CED vs. CRT	3.48	.87	-	-	2.92	1.03	85	4.54	.000	.59
CED vs. ID	3.48	.87	3.05	.99	-	-	83	2.81	.023	.47
<b>Usability</b>										
CED vs. CRT	3.38	.69	-	-	2.99	.73	85	5.04	.000	.56
ID vs. CRT	-	-	3.33	.70	2.99	.73	81	2.37	.009	.53
<b>Difficulty or Ease of Use</b>										
CED vs. CRT	3.75	.82	-	-	2.75	.92	81	8.95	.000	1.15
ID vs. CRT	-	-	3.78	.83	2.75	.92	77	5.99	.000	1.18
<b>Facilitate Productive Activity</b>										
CED vs. CRT	3.32	.80	-	-	2.99	.91	81	3.67	.010	.38
<b>Readability</b>										
CED vs. CRT	3.43	.95	-	-	3.07	.91	81	3.41	.021	.39

duced the greatest amount of questions and discussion about root causes, whereas the CRT produced the greatest amount of questions and discussion about process and construction.

**Discussion**

The CED was specifically perceived by participants as better than the CRT at identifying cause categories, facilitating productive problem solving activity, being easier to use, and more readable. The ID was easier to use, but was no different from any of the other tools in other aspects, except cause categories. Concurrently, none of the tools was perceived as being significantly better for causality or participa-

tion. Rather, the study suggested that the ID is another easy alternative for root cause analysis. The data also support that the ID leads groups to specific root causes in about the same amount time as the CED.

The results neither supported nor refuted the value of the CRT other than to verify that it is, in fact, a more complex method. Considering the effect sizes, participants defined usability as ease of use. Thus, the CRT was perceived by participants as too complex. However, usability was not related to finding root causes. Authors and scholars can reasonably argue that a basic requirement for a root cause analysis is the identification of root

causes. Until this can be statistically verified, one cannot definitively claim the superiority of one tool over another.

It is postulated that the difference in average process time was due to the construction complexity of the CRT. Also, the variability of the ID improved over time while the variability of the CRT deteriorated. If any conclusion can be reached from this, it is perhaps that as groups gain experience, the learning curve for the ID increases. Conversely, experience may be detrimental to the CRT because it uses a different logic system that starts with effects and moves to causes, whereas the ID starts with potential causes.

Specific observations indicated that more extensive facilitator intervention was required in the CRT process than on the other tools. Facilitators had to intervene more often because the groups had difficulty building the CRT without making construction errors or process mistakes. Most interestingly, in spite of difficulties, groups using the CRT were able to identify specific and reasonable root causes over half the time as compared to the CED and ID. This was one of the distinguishing findings of the study.

With the ID, groups were able to identify root causes, but only because ID construction requires a count of incoming and outgoing arrows. By simply looking at the factor with the most outgoing arrows, groups declared that a root cause had been found and stopped any further any critical analysis. Their certainty undermined critical analysis. Khaimovich (1999) encountered similar results in a behavioral study of problem solving teams where participants were reluctant to scrutinize the validity of their models and remained content with their original ideas.

So why were participants unable to recognize the difference between the three tools with respect to the integrity of their root causes? First, because the participants were asked only to identify a root cause, not a root cause that was reasonable and specific enough to be acted upon. Second, most of the groups avoided creating the tension that might have produced better results. To quote Senge (1990), "Great teams are not

characterized by an absence of conflict" (p. 249). Although the majority of the CRT groups were uncomfortable during the process, the quality of their outputs was better.

Third, groups were (a) learning the tools for the first time, (b) emotionally involved in the process, and (c) engaging in what Scholtes (1988) called the "rush to accomplishment." Because many participants were learning and doing, they did not have time to assess or reflect on the meaning of their outputs. Their reflection was impaired by the emotionally-laden group processes. In addition, participants were instructed to work on the problem until they had identified a root cause, which, in some groups, was manifested by the need to achieve that goal as quickly as possible.

### Implications for Policy and Practice

The type and amount of training needed for each tool varies. The CED and ID can be used with little formal training, but the CRT requires comprehensive instruction because of its logic system and complexity. However, the CED and ID both appear to need some type of supplemental instruction in critical evaluation and decision making methods. The CRT incorporates the evaluation system, but the CED and ID have no such mechanism and are highly dependent on the thoroughness of the group using them.

Serious practitioners should consider using facilitators during root cause analysis. Observations found that most groups had individuals who

dominated or led the process. When leadership took place, individual contributions were carefully considered with a mix of discussion and inquiry. Group leaders did not attempt to convince others of their superior expertise and conflicts were not considered threatening. In contrast, groups that were dominated did not encourage discussion and differences of opinion were viewed as disruptive. An experienced facilitator could encourage group members to raise difficult and potentially conflicting viewpoints so the best ideas would emerge.

These tools can be used to their greatest potential with repeated practice. Like other developed skills, the more groups use the tools, the better they become with them. For many participants in this study, this was the first time they had used a root cause analysis tool. Indeed, for some, it was their first experience with any structured group decision-making method. Their experience and participation created insights that could be transferred to other analysis activities.

### Conclusion

The intent of this research was to be able to identify the best tool for root cause analysis. This study was not able to identify that tool. However, knowledge was gained about the characteristics that make certain tools better under certain conditions. Using these tools, finding root causes seems to be related to how effectively groups can work together to test assumptions. The challenge of this study was how to capture and test what people say about

Table 6. Summary of Questions, Observations, and Tool Outputs

Tool	Output Evaluations			Questions and Observations			
	Technical Accuracy	Root Cause?	Integrity of Root Cause	Process	Root Cause	Construction	Group Dynamics
CED	1M, 8H	2Y, 7N	8L, 1H	17	11	11	17
ID	1M, 8H	8Y, 1N	3L, 4M, 2H	14	5	16	22
CRT	3L, 2M, 4H	8Y, 1N	2L, 2M, 5H	23	1	19	22

H = High, M = Medium, L = Low  
Y = Yes, N = No

the tools versus how the tools actually perform. Perhaps the answer lies in the interaction between the participants and the tool. Root cause analysis may be valuable because it has the potential of developing new ways of thinking.

**References**

Anderson, B., & Fagerhaug, T. (2000). *Root cause analysis: Simplified tools and techniques*. Milwaukee: ASQ Quality Press.

Anderson, N. H. (2001). *Empirical direction in design and analysis*. Mahwah, NJ: Erlbaum.

Arcaro, J. S. (1997). *TQM Facilitator's Guide*. Boca Raton, FL: St. Lucie Press.

Brassard, M., & Ritter, D. (1994). *The memory jogger II: A pocket guide of tools for continuous improvement and effective planning*. Salem, NH: GOAL/QPC.

Brassard, M. (1996). *The memory jogger plus+: Featuring the seven management and planning tools*. Salem, NH: Goal/QPC.

Brown, J. I. (1994). Root-cause analysis in commercial aircraft final assembly. (Master's project, California State University Dominguez Hills). *Master's Abstracts International*, 33 (6): 1901.

Cohen, J. (1988). *Statistical power and analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cox, J. F., III, & Spencer, M. S. (1998). *The constraints management handbook*. Boca Raton, FL: St. Lucie Press.

Dettmer, H. W. (1997). *Goldratt's theory of constraints*. Milwaukee: ASQC Press.

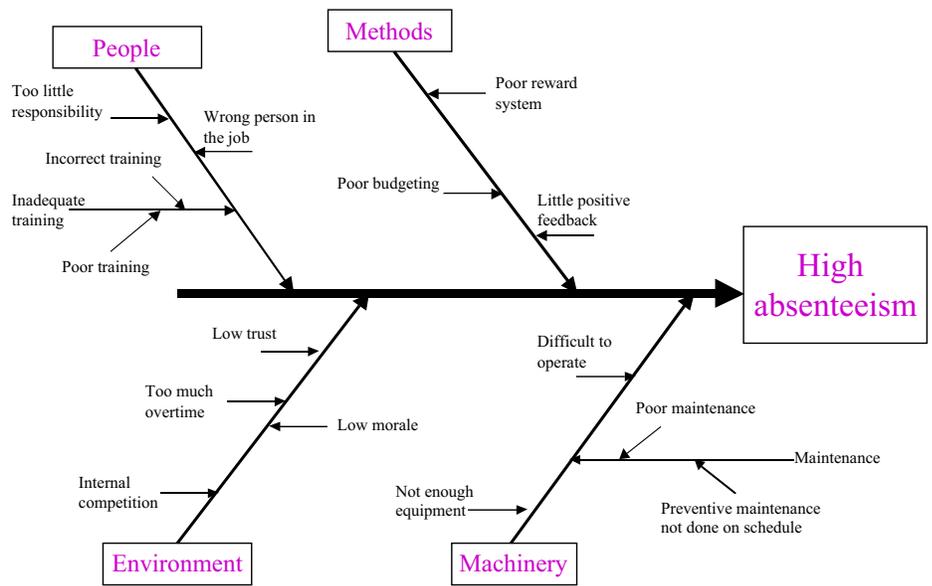
Dew, J. R. (1991). In search of the root cause. *Quality Progress*, 24 (3): 97-107.

Field, A. (2000). *Discovering statistics using SPSS for Windows*. London: Sage.

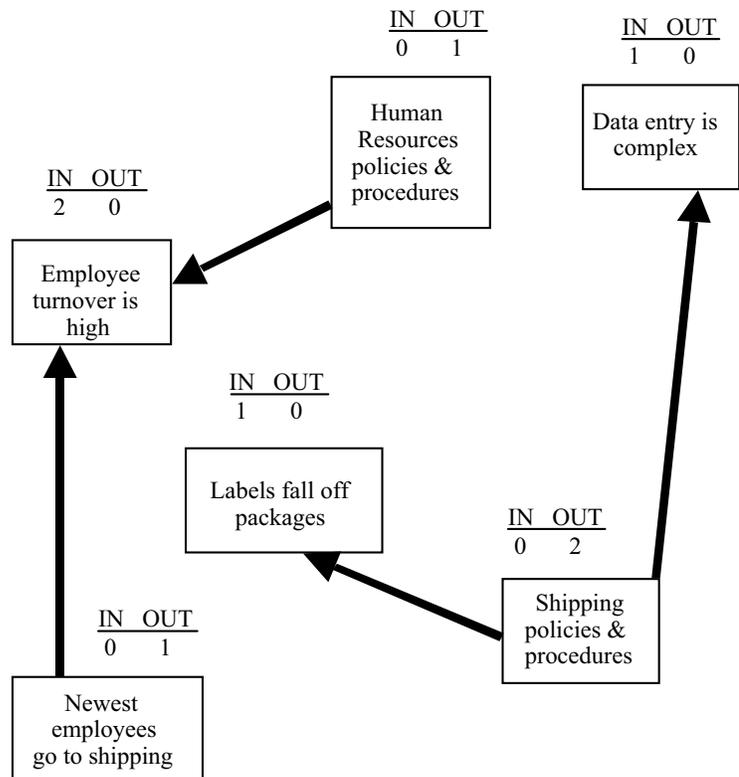
Frendall, L D., Patterson, J. W., Lenhartz, C., & Mitchell, B. C. (2002). What should be changed? *Quality Progress*, 35 (1): 50-59.

Girden, E. R. (1992). *ANOVA repeated measures*. Newbury Park, CA: Sage.

**Example of a Cause-and-Effect Diagram**



**Example of an Interrelationship Diagram**



Gliner, J. A., & Morgan, G. A., (2000). *Research methods in applied settings: An integrated approach to design and analysis*. Mahwah, NJ: Erlbaum.

Goldratt, E. M. (1994). *It's not luck*. Great Barrington, MA: North River Press.

Gravetter, F. J., & Wallnau, L. B. (1992). *Statistics for the behavioral sciences: A first course for students of psychology and education* (3<sup>rd</sup> ed.). St Paul, MN: West Publishing Co.

Huynh, H., & Feldt, L. (1976). Estimation of the Box correction for degrees of freedom from sample data in the randomized block and split plot designs. *Journal of Educational Statistics*, 1 (1): 69-82.

Ishikawa, K. (1982). *Guide to quality control* (2<sup>nd</sup> rev. ed.). Tokyo: Asian Productivity Organization.

Khaimovich, L. (1999). Toward a truly dynamic theory of problem-solving group effectiveness: Cognitive and emotional processes during the root cause analysis performed by a business process re-engineering team. (Dissertation, University of Pittsburgh). *Dissertation Abstracts International*, 60 (04B): 1915.

Lepore, D., & Cohen, O. (1999). *Deming and Goldratt: The theory of constraints and the system of profound knowledge*. Great Barrington, MA: North River Press.

Mizuno, S. (Ed.). (1988). *Management for quality improvement: The seven new QC tools*. Cambridge: Productivity Press. (Original work published in 1979).

Moran, J. W., Talbot, R. P., & Benson, R. M. (1990). *A guide to graphical problem-solving processes*. Milwaukee: ASQC Quality Press.

Pasquarella, M., Mitchell, B., & Suerken, K. (1997). A comparison on thinking processes and total quality management tools. In *1997 APICS constraints management proceedings: Make common sense a common practice*, Denver, CO, April 17-18, 1997. Falls Church, VA: APICS. 59-65.

Robson, M. (1993). *Problem solving in groups*. (2<sup>nd</sup> ed.). Brookfield, VT: Gower

Scheinkopf, L. J. (1999). *Thinking for a change: Putting the TOC thinking processes to use*. Boca Raton, FL: St. Lucie Press.

Scholtes, P. (1988). *The team handbook: How to use teams to improve quality*. Madison, WI: Joiner.

Senge, P. (1990). *The fifth discipline*. New York: Doubleday.

Smith, D. (2000). *The measurement nightmare: How the theory of constraints can resolve conflicting strategies, policies, and measures*. Boca Raton, FL: St. Lucie Press.

Snow, R. E. (1974). Designs for research on teaching. *Review of Educational Research*, 44 (3): 265-291.

Sproull, B. (2001). *Process problem solving: A guide for maintenance and operations teams*. Portland: Productivity Press.

Stevens, J. P. (1999). *Intermediate statistics: A modern approach*. Mahwah, NJ: Erlbaum.

Wilson, P. F., Dell, L. D., & Anderson, G. F. (1993). *Root cause analysis: A tool for total quality management*. Milwaukee: ASQC Quality Press.

Example of a Current Reality Tree

