Abstract
This paper examines the research question of how an initial problem definition can shift in scope from the start of a project to the end of a project. The initial problem statement came from a road construction experiment which would provide data to support the commercial use of Radio Frequency Identification (RFID) tracking and tracing technology attached to asphalt dump trucks. This field research experiment concluded in Summer 2010 in Anchorage, Alaska. This paper is a post mortem on the control and communications variables constraining the state of the problem statement within that experiment. This paper also indicates that variations from the initial problem definition may produce different experimental conclusions.

Introduction
The original research problem came to the researcher from a supervisor and later from a contract proposal from the Alaska Department of Transportation (AKDOT) to use RFID technology to track dump trucks in a closed loop transportation route. This RFID technology was a miniaturized radio transmitter mechanism (computer chip), connected to a thin antenna. This computer chip contained coded data for the truck identification and load identification. This data could be read by a receiver using an antenna that would be placed at various highway markers along a transport route that a dump truck would follow for this construction job. The data from the RFID tags would be read as the truck passed by the highway marker. The RFID tag data was also to be read at the asphalt supply site when the load was asphalt was placed in the truck bed and later, the data read as it arrived at the road construction site.

The current manual process of tracking was when a truck completed a delivery route with using the current paper and pencil recording method by personnel along the truck route. A AKDOT worker would use a pen or pencil to annotate on a piece of paper when the driver was at the different locations.

The use of RFID technology was seen as a way to more accurately record those times compared to the manual method. It seems that AKDOT history showed that there were significant errors or disputes regarding these times were accurately reported by workers. So, a simple comparison of a routine construction site operations using manual data recording was to be compared with an automated system to document any frequency of data reporting errors. (Hedgepeth, 2010)

The RFID Technology
The RFID technology is composed of a small plastic hard shell about five inches long, containing a microchip with a coiled antenna. This RFID tag is fixed at the rear of a construction truck, usually near the license plate. The black rectangular item to the left of the license plate in Exhibit 1 shows an RFID tag.

Exhibit 1. RFID tag, dark plastic rectangle, placed to the left of the license plate. (Hedgepeth, 2010)

This RFID tag produces an electronic signal that can be read by a radio and antenna system placed usually with a five to ten feet of the rear of the truck. (Hedgepeth, 2007)

The use of such RFID technology for commercial applications in the road construction business has been confined to pilot tests and laboratory tests from universities. This was the first time that RFID technology had been used in an Alaska road construction project. (Hedgepeth, 2010)

The Problem
A statement from the original contract was...

"...that there was no real-time inventory capture of data. There was no real-time communications of this data capture. The current process is manually receiving and recording carrier and shipping and load data.” (Hedgepeth and Henrie, 2011, 2)

This paper will demonstrate how field implementation can create new variables as well as new assumptions and invalidate initial project assumption. To
state a problem, you must have some idea of what to measure, some metric. This original problem statement defines a method to analyze the variables within this problem which is a measure of location variables using distance and time.

The Assumptions
The assumptions from the original contract were created from a team of AKDOT engineers and the researcher. These assumptions included:

- AKDOT field data is captured with a normal probability distribution.
- Errors in one AKDOT database are independent of errors in any other database.
- AKDOT manual and automated data and files are not damaged or altered prior to study team having accessibility.
- The data capture technology for RFID records do not fail and event records are uniquely identified.
- The probability of each RFID and manual data record being correct in any database is the same.
- The AKDOT and other construction professionals do not make mistakes when checking or providing the study team data or access to trucks, pavers or other equipment. (Hedgepeth and Henrie, 2011, 3)

The assumptions further define the problem and the possible metrics in statistical terms, balanced with the potential for technical or human error in any data collection.

The Metrics
While the problem statement identified a limited metric of truck movement over time, plus some variable of transmitting this data from one source to another, the metrics from the original contract were changed after the contract was let to the following:

- The number of truckloads delivered in 24 hours.
- The number of human errors in data capture.
- The cost of transportation to include fuel to the carrier.
- The amount of time driver is behind the wheel. (Hedgepeth, 2010).

During the process of early field trials and testing of the tracking technology the following metrics were identified to be captured:

- Tracking data accuracy
- Asphalt loading time per truck
- Truck transportation time
- Truck and driver waiting time
- Truck unloading time for asphalt
- Truck return time for a new load of asphalt
- Truck waiting time to load asphalt mass
- Truck driver signature capture time
- Asphalt temperature during loading and unloading
  (Hedgepeth and Henrie, 2011, 4).

During the field data collection process the data recording devices were programmed to the following metrics:

- Ticket number
- Customer name
- Truck identification
- Net weight of truck
- Asphalt plant name
- Truck arrival date and time at plant
- Truck loading date and time
- Plant departure date and time
- Time truck at loading plant
- Paver identification
- Truck job arrival date and time detected by paver
- Paver arrival date and time
- Paver departure date and time
- Truck job left date and time by paver
- Total time on job for truck
- Round trip time for truck
- Job ID on the ticket
- GPS latitude coordinate of the paver
- GPS longitude coordinate of the paver

The final metrics used were:

- Time of travel of the trucks
- Distance driven by the trucks
  (Hedgepeth, 2010)

The number of variables to be used to collect data from the beginning of the experiment to the end is graphically show in Exhibit 2.

Exhibit 2. The flow of the number of variables to be used to measure the truck movement during the course of the experiment.
In the beginning there were just two variables, then four, then nine, which increased to 19 due to the computer software data base, and finally back to two variables. The final two variables contained only one common variable with the initial two variables, the time of travel by the trucks.

Before the experiment began, the contract called for tracking and tracing the route of only one dump truck. During the initial first days of the experiment the contractor modified to track all ten trucks being used in the road construction project. This increased the expenses of the project and complexity of RFID placement on the trucks and discussion of different locations of radio collection points along a variety of possible routes.

The Data
The raw data collected from the original contract is shown in Exhibit 3. The total trip data is recorded in hours and number of trips made from the 10 trucks used.

Exhibit 3. Raw data from Alaska RFID research

This data indicates an expected time of about one hour for a complete round trip per truck based on more than 100 such trips. There were data collection errors, both technical recording and manual input, that are listed as a zero time value. The spikes shown are valid and confirmed data entries where trucks made the round trip in two, five and nearly six hours. And, all assumptions were violated.

Definition of Problem
A problem can be defined as a concise statement describing a complex entanglement of controlled human viewpoints and communications balanced in a robust environment that evolves into an unstable component of a system.

Kerlinger (1986) provided a simpler definition. He states, “the problem should express a relation between two or more variables” (p. 16). In its simplistic and almost cause-and-effect approach to this term, he is stressing the relationship between (at least) two variables.

A good way to express the problem is through a question. For example, is sunshine related to plant growth? You can see that you have two variables: sunshine is one variable and plant growth is the second variable. We could add a constraint to narrow the focus of this problem. What if the geographic area is added? Is sunshine related to plant growth in Anchorage, Alaska? What assumptions are you making with this question? If you know that Anchorage, Alaska has a temperature range that goes from -40 to 60 degree F, does this time factor impact the focus of the question, the focus of the problem? If the plant is a tomato plant the temperature will impact the growth. If the plant is a blue spruce tree, the temperature will not impact the growth. So, the type of plant becomes another constraint factor. The question might need to be reworded to include additional constraints. Use of the question format of the problem statement can simplify the complexity of these constraints. The “questions have the virtue of posing problems directly” (Kerlinger,1986, p. 17). But, how do you pose a question that will include those constraints that are important for academic scholarly research?

Problems that pose a relationship between two variables, A and B or sunshine and plant growth, need to have a consistent metric. This metric is the data that is collected that defines variables A, B, sunshine, and plant growth. The sunshine could be described as a data item of yes or no, a binary number. The yes could be the number symbol 1. The no could be the number symbol 0. The 1 or yes represents the presence of sunshine. The 0 or no represents the absence of sunshine; we could call it night time. Then, we could redefine the 1 as daytime and the 0 as nighttime. A number symbol of 1 could indicate plant growth. The symbol 0 could represent the absence or growth; notice we did not indicate death of the plant, just growth!

The equation of this question could be written in a heuristic form as shown in the word equations (1) and (2) or (3) below:

\[
\begin{align*}
\text{If sunshine} = 1, \text{plant growth} & = 1 \quad (1) \\
\text{If sunshine} = 0, \text{plant growth} & = 0 \quad (2) \\
\text{If sunshine} = 0, \text{and location} & = \text{Alaska, and plant} = \text{tomato, and temperature} = -40, \\
& \text{then plant growth} = 0 \quad (3)
\end{align*}
\]

The word equation could be written as a math model as:

\[
\text{Plant growth} = f (\text{sunshine, location, plant type, temperature})
\]
In this case study, the variables seem to be more than two. But, what is also missing is a question. What is the problem statement indicating toward the inventory tracking over time for trucks delivering asphalt to a road construction site? A possible question could be that using RFID tracking technology is more efficient compared to manually tracking truckloads of asphalt. Such a question might be, “How efficient is manually tracking data compared to an automated tracking of data?” This problem statement or question seems appropriate, but is it?

Kerlinger (1986) indicates that defining a problem is a complex association of “psychological, sociological, and educational reality” (p. 23).

We will not delve into the aspects of these three multivariate (Kerlinger, 1986) factors, but the wise PhD student should consider the implications of such factors when formulating their dissertation problem. For a contractor publishing a request of a research proposal, these may not enter into their mind as the contract purpose may seem clear. A contractor may see a problem as a simple question such as, if we add RFID technology to dump trucks we can decrease delivery times. A contractor in state transportation may not see the value of the complexities of the constraints around the term problem as shown in Exhibit 4, or be concerned with any psychological or sociological aspect or educational reality of dump truck drivers, researchers, and the weather in Anchorage, Alaska.

Klaus Krippendorff (1986) defines a problem as the following:

Literally, something thrown forward (in time). Specifically, a cognitive inSTABILITY of disposition which demands of an organism that something be done to change its current BEHAVIOR or [take] it away from an existing undesirous and problematic STATE. Problems need not be caused by disturbances from the organisms environment (adaptation) but may also be the product of cognition itself, particularly when the brain is complex and self-organizing (SELF-ORGANIZATION, PROBLEM-FORMULATION, PROBLEM-SOLVING) (p. 60).

Krippendorff defines his terms with links to other related terms, those listed in all capitalized letters. When one maps the definition of the term problem, one can develop a structure as shown in Exhibit 4. Each term on the right side is linked to the seed term we call a problem. We can view this mapping or tree structure of the terminology as a first level of concepts or other terms that constraint the term problem.

Exhibit 4. A word mapping of terms that constrains the definition of the term problem (Krippendorff, 1986).

<table>
<thead>
<tr>
<th>PROBLEM = ∑</th>
</tr>
</thead>
<tbody>
<tr>
<td>STATE</td>
</tr>
<tr>
<td>BEHAVIOR</td>
</tr>
<tr>
<td>SELF-ORGANIZATION</td>
</tr>
<tr>
<td>inSTABILITY</td>
</tr>
<tr>
<td>ADAPTATION</td>
</tr>
<tr>
<td>PROBLEM-FORMULATION</td>
</tr>
<tr>
<td>PROBLEM-SOLVING</td>
</tr>
</tbody>
</table>

The problem is then several components describing a complex entanglement of psychological, sociological, and educational realities as indicated by Krippendorff (1986).

Given the initial problem and assumed metrics, this reality is then part of human viewpoints and the human and technology-based communications process. Add a robust environment of uncontrollable factors from field exercises, and one can be left with a problem statement that will evolve into an unstable definition.

Evaluation of the Problem

The initial contract called for using one truck. During the field trial set up, this number changed from one truck to 10 belly dump trucks. (Hedgepeth, 2010)

The metrics changed during the course of this project, from the discussion stage of a statement of work to the placement of the tracking technology in the field, identifying two variables up to 19 and then back to two different variables from the original two. The degree of changes to the metrics expanded to include geographic details. The expansion of detailed data collection to include truck and paver identification and arrival and leaving times, were part of a preset technology data capture procedure. Previous experience from Minds, Inc., which provided the tracking hardware for data collection, had previous experience. This previous experience included more detailed data than was needed to address the scope of the problem. While not totally a data overload problem, there was an issue of dirty data being collected. The data provided an inconsistent picture of the tracking and tracing needs of the trucks; it was fragmented with missing data points that went to zero values (Davenport and Harris, 2007; Hedgepeth, 2007). When the metrics change, it changes the p r o b l e m s p a c e , changes the problem statement.

Metrics are the skeletal framework of a research problem statement. The significant shift in metrics along the route of a field exercise that was complete in 30 days, is problematic.
To minimize the problem statement from shifting due to improper or inconsistent metrics of variables to capture, the following steps should be followed from Hedgepeth (2007):

- Listen to the customer. The customer or decision maker is always struggling to understand their problem. They have not only internal metrics but there are external metrics. The external metrics are part of the environment that may not be in total control of the customer or the researcher.
- Define the problem. This is usually the first step in many books on operations research. One of the problems in this case study was the technology solution that was specified by the customer without clear understanding of underlying assumptions and potential external metrics of data capture.
- Assemble the data. In the use of RFID, “there is no data on how the new RFID system will perform.” (Hedgepeth, 2007, p. 37).
- Formulate a view of the problem as a model or flow chart of other visual aids. Look for a pattern in possible variables that might be needed to define the problem space, to meet the objectives of the customer. This involves knowing when and how to collect that data. “Rarely are data trends or patterns so obvious that you can make a new discovery from them.” (Hedgepeth, 2007, p. 38).
- Develop a model. This is a basic math model to a simulation, visual computer model. What you are looking for are the unknown unknowns (Hedgepeth, 2007, p. 39) that come from those psychological, sociological, and educational aspects mentioned by Kerlinger (1986). It was these unknown unknowns that became knowable too fast within the 30 days of this experiment, this field research project.

At this juncture the feedback of all these steps into a problem space should provide a more robust problem statement, less fragile in metrical skeletal framework.

Conclusion
The problem in this case study was simply defined, but not concise in meaning or scope. The complexity of the metrics changing over time did not surface. The complex entanglement of human viewpoints contributed to the emergence of unknown unknowns. The technological metrics from the data gathering hardware and software lead to potential data overload, but resulted in contributing to dirty data being collected. The communications between the automated data collection and the manual data collection of past business construction operations lead to corrupted data. All assumptions failed. The end result was the evolution of an unstable system of data tracking and tracing from what seems like a simple exercise.

Future Research
Future research such as this case study should adhere to a scholarly approach using measurable problem statements and hypotheses such as the following:

Problem Statement: How are the asphalt truck ticket capture and filling accuracy rates affected by the introduction of RFID tracking system? (Hedgepeth, 2010, p. 11).

Hypothesis One (H1): The current manual (paper based) truck ticket tracking accuracy (MTA) rate is equal to the RFID tracking accuracy (RTA) rate (Hedgepeth, 2010, p.11)

\[ H_1: MTA = RTA \] (5)

During any field test, the researcher would be stating with equation (5) with the above hypothesis that it did not matter if some RFID tracking method was used compared to a manual method of data collection. The results would be expected to be not be significant. Proving significance would indicate that one data collection method is more accurate to the other however.

Hypothesis Two (H2): The manually processed truck ticket accuracy rate demographic (MRAD) for two-level factors are the same (Hedgepeth, 2010, p. 11).

\[ H_2: MRAD_1 = MRAD_2 \] (6)

During any field test, the researcher would be stating with equation (6) that the number of error collected between the two methods would not be significant. However, if there was a statistical significance in the number of errors, then one method would be considered possibly more accurate in data collection.

This problem statement and hypotheses include the potential for the demographics of the different types of drivers, and all employees involved in this process (Hedgepeth, 2010, p. 11).

Acknowledgements
This Alaska research project reported in this study was funded by AKDOTPF/AUTC Project No. C-08-01. The field data collection was undertaken by RFIDcomplete, LLC. MINDS, Inc. provided technical assistance, hardware and software. Research grant contract administration was through the University of Alaska Anchorage (UAA) College of Business and Public Policy (CBPP).
References
Kieres, Curtis, University of Alaska – PaveTag Summary, MINDS, Inc. (2010)
Krippendorff, Klaus, A Dictionary of Cybernetics, University of Pennsylvania, (February 1986).

About the Author
Dr. Oliver Hedgepeth is Program Director of Reverse Logistics Department, School of Management, American Public University System. His Ph.D. is in Engineering Management from Old Dominion University. Research interests are in reverse logistics and project management. His book, RFID Metrics, was published in 2007 by CRC Press.