

Quality and Decision Variables: Impact-Focused Taxonomy

Zbigniew J. Gackowski, California State University Stanislaus, 801 W. Monte Vista Avenue, Turlock, CA 95382; E-mail: ZGackowski@csustan.edu

ABSTRACT

This paper presents an inquiry into the applicability of operation qualities identified for data and information to decision variables in general. It posits and demonstrates that they apply equally to data values, information values, elements of knowledge, physical factors in operations, and thus to decision variables used in decision-making to represent them. This approach leads to a universal hierarchical impact-focused taxonomy of decision variables of theoretical and practical importance in analyzing decision situations, their models, and in prioritizing research on decision situations.

Keywords: Quality, data, information, rules of reasoning and proceeding, operation factors, decision variables, universal taxonomy

INTRODUCTION

The theoretical and empirical research on data and information quality conducted for more than a decade and the operations-research approach to quality of factors in operations elucidate that the identified principles apply equally to decision variables in decision-making. On must, however, take a broader than the internal ontological view on quality. The teleological pragmatic viewpoint common in operations management, operations research, management sciences and decision sciences elucidates that the identified principles of operation quality apply to data values, information values, elements of knowledge (rules of reasoning and proceeding), physical factors in operations, and thus to the decision variables, which represent them all.

The presented approach is a purely theoretical one. Consequently, the presented taxonomy claims universal theoretical and rigorous practical validity. Survey-based empirical approaches such as TQM and TDQM are not considered here. They are of high practical value but alas of low scientific validity as they always are situation-specific only.

The main contribution of this paper is a universal hierarchical impact-focused taxonomy of decision variables. The main purpose of this paper is to present this concept for challenge and discussion. For focused reading, key terms in paragraphs are in **bold** font, emphasis in *italics*, highest emphasis underlined, and terms followed by a definition are in **bold italics**.

A SHORT REVIEW OF LITERATURE

Based on **ontological foundations**, Wand and Wang (1996) proposed four data quality dimensions (complete, unambiguous, meaningful, and correct) that are intrinsic to system design and operations. However, they labeled them intrinsic to data. At the same time, Wang and Strong (1996) using empirical approach identified other about 179+ (later reduced to about 15-20) dimensions of quality mentioned in questionnaires by individuals representing data users or consumers.

Liu and Chi (2002) categorized different approaches to data quality as intuitive, empirical, and theoretical. They concluded, the “*Existing theoretical approaches are limited in their ability to derive a full-fledged measurement model*” and a “*generally accepted model has not yet appeared*” (p. 292). They developed a concept of **evolutional and theory-specific data/information quality** that evolves along the stages of data collection, organization, presentation, and application.

Oliviera, et al. (2005) claimed they presented “**A formal definition of data quality problems.**” The paper identifies and defines 30 specific possible distortions inflicted onto a set of entered data values assumed to be correct and updated. However, their formal definition is limited only to data in databases.

Anchoring the concept of data/information quality in operations research, management science, and decision science, Gackowski (2005) defined a theoretical teleological content-focused framework of **operation quality requirements** of data and information values viewed from the perspective of operations management. They pertain to data, information, elements of knowledge (rules of reasoning and proceeding), and any factor of substance.

QUALITY AND OPERATIONS

Point and Frame of Reference, Observer, Assumptions and Postulate

There are two fundamental keywords, ‘**quality**’ and its adjective ‘**operation**’ because it pertains to quality of factors in operations. A reliable theory must refer to a well-defined point of reference, observer, frame of reference, and a yardstick for measuring the results of operations:

Assumption 1a: The **main purpose** of operations is the main **point of reference** and it is measurable.

Assumption 1b: The **observer** is the decision-maker. Decision makers do not act in the actual reality – a **subject of science and technology**, but in their subjective reality, which is the **subject of phenomenology**. Usually, the two realities overlap only partially: (a) completely, only in structured decision situations, (b) not at all, in fully irrational decision making, and (c) partially, in semi structured decision situations, when part of the model is scientifically and technically sound, and the rest is subjective of unknown validity, hence of unpredictable outcome. Thus, decision situations are complex structures, which might be better handled by **complex analysis** (as defined in mathematics with complex variables, which are pairs of real and imaginary components) and become an extension of the theory of decision-making.

Assumption 1c: The **frame of reference** consists of:

- **sn** – a vector of all **states of nature** – independent environmental variables, which are beyond control of decision-makers and of significant materiality or impact on operations; they describe a part of the **circumstances** under which operations occur
- **V** – set of significant (by impact) dependant variables **v** under decision-makers’ control,
- An adopted **criterion of effectiveness** of operations, and
- **Assumption 2:** Decision-makers employ only **rational** and **rule following** choices as defined by March (1994).

As long all of the above including the assumptions do not change, each variable:

- Is viewed and assessed the same way by rational decision-makers (observers),
- Is bound by the same logic, and
- All principles, to which qualities of operation factors are subject, remain unchanged.

Postulate of general relativity of variables: Rational observers view, perceive, and assess the same way all variables, as long the **purpose** and **frame of reference** remain unchanged. Nevertheless, when changes occur, it changes how even objectively identical operation factors and their qualities are seen, perceived, observed, and assessed by decision-makers. This is the **all-pervasive principle of relativity of all variables in informing and in all operations** in general. It employs the analogy of inertial frames in theoretical physics.

A unit of any adopted **measure of results** of operations M_{RO} may serve as a **yardstick**. This is under the assumption that M_{RO} is a function of the main purpose P , all states $s(v) \in S(v)$ of significant variables $v \in V$, and of all significant states of nature denoted by vector sn , formally

$$M_{RO} = M_{RO}[P, s(v), sn] \text{ for all } s(v) \in S(v) \text{ and } v \in V$$

(Assumption 3)

For instance in business, income after taxes, retained earnings, return on investment, return on equity, cost effectiveness, etc may measure the results. In public administration, measurable or only observable results can be derived from the entity's mission. In military operations, they may be described by the expected tactical or strategic objectives. For example, when cost effectiveness $C_E(O)$ of operations O matters, then the percentage point of the ratio of the main purpose, P divided by the cost $C(O)$ of operations O over time may serve as a unit of measure. Then formally: $M_{RO} = C_E(O) = 100 * P / C(O)$.

All of the above imply that a relatively complete qualitative cause/effect diagram of operations (known also as a fishbone diagram) is available or can be drawn. Such a diagram tries to identify all factors of significant materiality in the situation, the required actions to implement the decisions made, and/or the results.

Operation quality is defined here by distinct significant states of the factors' qualities (attributes, dimensions) that enable them to play a significant role in operations. If so, they all must be represented by corresponding variables in the decision situation matrix.

These factors can be **factors of substance** and **factors of symbolic nature**. **Factors of substance** may be the four M's (material, method, machinery, manpower) and others including respective states of their significant qualities. In contrast to factors of substance, **factors of symbolic nature** may be data values, information values, and elements of knowledge meant as rules of reasoning and proceeding, again including respective states of their significant qualities. Operations factors may be:

- Already **available** such as any available substance, data, or elements of knowledge, and
- **Not yet available**, to be acquired or as yet even unknown to be recognized – such as any additional substance, information, or element of knowledge.

Operation variables acquire potential **materiality** from the purpose, circumstances and by the adopted criterion of effectiveness of operations. **Materiality $M(v_x)$ of a specific variable v_x** is defined as the difference in results of operations RO using M_{RO} as their measure when acting with all variables V and without the specific variable v_x . Formally:

$$M(v_x) = M_{RO}(V) - M_{RO}(V - v_x)$$

Significant materiality or impact on operations is defined by the condition $M(v_x) \geq \text{Min}(\Delta M_{RO})$, where $\text{Min}(\Delta M_{RO})$ is the threshold of significance expressed as a minimal increment ΔM_{RO} of the measure of results M_{RO} the decision-makers care. Each decision variable v belonging to their set V must meet the following condition. Formally:

$$V = E [M(v) \geq \text{Min}(\Delta M_{RO})] \text{ for all } v \in V$$

After any new variable ΔV and its state ΔS has been recognized as significant, the previous set of significant variables V_p and their states S_p will be updated (augmented or reduced) to their respective current states S_c of variables V_c . Formally:

$$V_c = V_p + \Delta V \text{ and } S_c = S_p + \Delta S$$

In general, **quality requirements $QR(v)$** for variables $v \in V$ can be represented as a vector of required states of their quality $rs(q(v)) \in RS(q(v))$, and all of them can be represented by corresponding decision variables. Formally:

$$QR(v) = [rs(q_1(v)), rs(q_2(v)), \dots rs(q_n(v))] \text{ for all } q(v) \in Q(V), v \in V$$

Universal Taxonomy of Decision Variables

One of the first steps to knowledge is clear distinction of entities (objects, qualities, states, relationships, phenomena, etc) under investigation. One needs a rigorous taxonomy of qualities, their states, and the consequences of their changes. In science, the strongest taxonomy is the binary one. It is natural and logically perfect in accordance to the fundamental principles of thinking.

Most textbooks and empirical studies list under different names a plethora of qualities or dimensions of data/information quality for consideration as stated by Wang and Strong (1996). In operations, certainly one may identify easily hundreds or thousands more qualities. The major question is, however, how to examine those qualities in real life situations, how to focus the attention of the examiners, how to provide them with diagnostic guidelines? Which of them affect the situation results directly or indirectly, which are primary or secondary, necessary or optional, which are not fully attainable and therefore one must learn to act with only some acceptable level of quality. Because many are there, these qualities require a systematic uniform approach to research and their practical diagnostic examination.

In informing, these concerns lead to a universal taxonomy of all known and not-yet known data or information qualities and requirements related to them, which together with research in-progress on interdependencies among them (Gackowski, 2004 and 2006) provides many clues in this regard. The same applies to their taxonomies. Thus by the power of this abstraction one can move from the universal hierarchical impact-focused taxonomy of quality requirements for data and information values (Gackowski, 2005) to an equivalent taxonomy of decision variables in decision-making (see Table 1), which does not yet cover the distinction of their real and imaginary components.

Operation quality of data of information (Gackowski, 2005) identified only **five** direct universally necessary quality requirements, **four** direct secondary quality requirements, when economy of operations matters, and some situation-specific ones, which make them usable. Hence, a variable uv is usable, if:

$$uv \equiv (\text{Interpretable} \vee \text{Recognizable}) \wedge \text{Relevant} \wedge \text{Significantly material} \wedge \text{Operationally timely available} \wedge (\text{Actionably credible} \vee \text{Actionably reliable}) \wedge \text{meets all situation-specific necessary quality requirements.}$$

Similarly, an **economically usable variable – euV** must be first usable (uv – see above) and then meet the four direct secondary requirements, formally:

$$euV \equiv uv \wedge \text{economically [(interpretable} \vee \text{recognizable)} \wedge \text{available} \wedge (\text{credible} \vee \text{reliable})].$$

To facilitate its comprehension it is necessary to reemphasize that in decision-making a variable (whether dependent or independent, deterministic or stochastic, discrete or continues) must represent all its significant states.

The **universal hierarchical impact-focused taxonomy of sets of significant variables:**

Table 1. Schema of hierarchical impact-focused taxonomy of decision variables

Categories of Decision Variables				
Direct			Indirect	
Primary		Secondary		
Universal	Situatio
y	n-
Necessary	Specific		
.....			

1. The taxonomy subdivides the universe of **variables V** into **direct** and **indirect** or subordinate variables.
 - a. A change from the previous state s_p to the current state s_c of **direct variables s(dv)**, where $dv \in DV \subset V$ immediately affects the decision situation itself, and/or the actions to implement the decisions made, and/or the results of operations, which implies they change the value of the adopted measure of results of operations ΔM_{RO} , formally: $(s_p(dv) \neq s_c(dv)) \Rightarrow (\Delta M_{RO} \neq 0)$. If any of the above listed requirements cannot be met, this implication remains valid.
 - b. A similar change of states of **indirect variables s(iv)**, where $iv \in IV \subset V$, as the name suggests, only indirectly affects the situation, for it determines or contributes to states of other **indirect variables of a higher order** (closer to the direct ones and at the extreme to **direct variables**). When s_p and s_c respectively denote **previous state** **current state** of an indirect or direct variables, and iv_n and iv_{n-1} respectively denote **indirect variables** of n^{th} order and **indirect variables** of higher $(n-1)^{\text{th}}$ order, for $n = 1$ indirect variable of 0^{th} order is a direct variable $iv_0 = dv$. It implies that changes of states of indirect variables of n^{th} -order causes a change of state of the related indirect variables of a higher order iv_{n-1} , or at the extreme of direct variables. Formally: $(s_p(iv_n) \neq s_c(iv_n)) \Rightarrow (s_p(iv_{n-1}) \neq s_c(iv_{n-1}))$. Example: If any of the twenty indirect qualities identified by Gackowski (2006a) as contributing to “actionable credibility” among them definition, variability, bias correctness, precision, and currency cannot be met at least at an acceptable level the concerned variable must be dropped from the decision situation matrix and its definition at least partially redefined.
2. The **direct variables** are subdivided into **direct primary** and **direct secondary** variables. The primary ones are Boolean {true, false}, that is exists or not, a requirement is either met or not.
 - a. Changes of states of the **direct primary variables s(dpv)**, where $dpv \in DPV \subset V$ result always in *qualitative* changes to the decision situations under consideration, which result in adding or eliminating a variable from consideration - labeled ΔV . Such changes add or delete entire rows or columns from the matrix that represent the model of a decision situation. When the above requirements of usability cannot be met, again, it requires at least a partial redefinition of the decision situation, which leads to quantitative consequences, as well. Formally: $(s_p(dpv) \neq s_c(dpv)) \Rightarrow [(V_p \neq V_c) \wedge (\Delta M_{RO} \neq 0)]$, where $V_c = V_p + \Delta V$.
 - b. Changes to states of the **direct secondary variables s(dsv)**, where $dsv \in DSV \subset V$ mainly *quantitatively* change the results of operations; hence, they may not necessarily be of significance ($\Delta M_{RO} \geq \text{Min}(\Delta M_{RO})$). Nevertheless, if the subsequent quantitative changes reach a critical point that is, if the current state $s_c \in C(s(dsv))$ belongs to the set of critical states **C**, they may trigger a qualitative change of situations. Then they become necessary, as well. The secondary variables are mostly of economic nature. If not only effectiveness, but also economy of results matters, the secondary variables also may become necessary, however not universally necessary. Formally: $(s_p(dsv) \neq s_c(dsv)) \Rightarrow [(\Delta M_{RO} \neq 0) \wedge \text{If}(s_c(dsv) \in C(s(dsv)) \text{ then also } (V_p \neq V_c))]$. Example: Usually environmental requirements must be met within a certain prescribed range, of course at cost. However, when the upper lawfully acceptable level has been exceeded, heavy fines may be imposed and even the entire operation suspended.
3. The **direct primary variables** are divided into those of **universal necessity** versus those that are necessary in specific situations, **situation-specific necessary**. Changes to their states are Boolean {true, false} and always redefine the decision situation by adding or deleting entire rows or columns from the decision situation matrix.
 - a. The **direct universal primary variables** are always necessary. Changes to their states $s(dupv)$, where $dupv \in DUPV \subset V$ add or eliminate them from consideration. Formally: $(s_p(dupv) \neq s_c(dupv)) \Rightarrow (V_p \neq V_c)$ always redefines the decision situation.
 - b. The **direct primary situation-specific variables** are necessary also. Changes to their states $s(dpssv)$, where $dpssv \in DPSSV \subset V$ also add to or eliminate variables from considerations, however they are not universally necessary only under situation-specific conditions. Formally: **If situation requires** then $(s_p(dpssv) \neq s_c(dpssv)) \Rightarrow (V_p \neq V_c)$. Example: Restricted availability of information in a competitive situation, otherwise, when available to all, it may lose its advantage, hence materiality).

Research Priorities

Once a universal hierarchical result-determined taxonomy of variables has been defined, it nearly automatically prioritizes the sequence of their diagnostic examination and any corresponding research. These priorities, again, do not yet cover the distinction of the real and imaginary components of variables.

1. **First priority.** The variables that represent the **direct universal primary requirements** should be the immediate subject to scrutiny. By the principle of usability (Gackowski, 2005), they must be **interpretable** or **recognizable**, **relevant**, of **significant materiality** of impact, operationally **timely available**, and **actionably credible** or **reliable**. If any of the five requirements is not met, the variable that represents the affected decision factor is out of commission. In addition, in operations, any factor must be **effectively complete** that is be usable together with other usable factors for a task to be accomplished, which gives us the sixth universally necessary requirement. All the above suffice for only effective operations, not necessarily economically effective. It pertains to all-out efforts such as special operations or terror acts, where economy is secondary. All the above requirements, if not met, call for a qualitative redefinition of decision situations. Materiality measures the importance of the variable.
2. **Second priority.** When economy matters, which is an all-pervasive principle in business, not necessarily in administration or military operations, four additional **direct secondary requirements** play an important role. Then the affected factor (represented by a decision variable) must be **economically interpretable** or **recognizable**, **timely available**, **actionably credible** or **reliable**, and **economically effectively complete** together with its companion factors. Now, the direct secondary requirements are also necessary, however, not universally necessary. Any changes to them, usually cause quantitative changes only, as long, any of them does not reach a critical state (for example, cannot exceed the limits required by law). Then again, it will cause a redefinition of the decision situation.
3. **Third priority.** All the rest belongs to **indirect quality requirements**. Depending on the length of the chain of their interdependencies, there are indirect requirements of the first, second, and subsequent orders. This fact again, clearly prioritizes the diagnostic sequence of their examination and respectively any related research about these factors.

Components of the proposed framework and model for research are anchored in basic scientific principles. Hence, they do not require extensive empirical validation except for coming up with examples to the contrary or other objections with regard to the logic of the model. Then the model may require a revision. The proposed model and framework needs, however, be discussed and challenged.

SUMMARY, CONCLUSIONS, AND LIMITATIONS

A formal definition of operation quality requirements of data and information, and any other factor in operations offers an insight that with no or only some terminological modifications it equally applies to factors of substance. In decision-making, all of them, if only significant, are decision variables. They all are decision variables subject to the same universal hierarchical impact-focused taxonomy. If this taxonomy remains substantially unchallenged, it qualifies as basic research in contrast to applied research of situation-specific limited validity.

Some principles of operation quality by which the quality requirements are governed, and pertain to decision variables may be summarized as follows:

1. **The principle of relativity of variables in decision-making.** Quality requirements are determined by the purpose **P** and circumstances of operations described by the vector **sn** of significant states of nature, the criterion of assessment of effectiveness of operations, and the assumption that decision-making is limited to rational and rule following choices (see the Postulate of Relativity).
2. **The principle of pervasiveness of materiality of factors.** A factor in operations confers its materiality upon all its qualities and its corresponding necessary task-specific usable companions represented by corresponding decision variables.
3. **The principle of usability of factors in operations** (when it meets all the **universally necessary (interpretable or recognizable, of significant materiality, operationally timely available, and actionably credible or actionably reliable)** and the other **situation-specific necessary quality requirements** that pertain to the corresponding decision variables.

4. **The principle of degradation of decision situations by declining usability of factors represented by variables.** If the usability of a factor:
- Is certain**, the decision-maker deals with a **deterministic situation** at least in the area affected by the factor
 - Is only probable** (the most likely case), the decision-maker deals with a **stochastic situation** to the same extent as above
 - Is not attained**, for instance when not timely available or not actionable reliable/credible, the decision-maker **games** to the same extent. It may be the case even when operations are not triggered, for instance when a threat is ignored.

This first attempt assumes a single main purpose with no conflicting requirements and constraints imposed upon decision makers. The impact-focused taxonomy of decision variables opens the door for research by modeling and simulating decision situations, as systems of state transitions, where the results by whatever adopted measure are a function of states of decision variables. Simulations will facilitate the quest for more complex quantitative dependencies and likely the discovery of other yet unknown interdependencies. Based on a formal theoretical model, research results of simulations conducted under rigorously controlled conditions promise results of a lasting validity when compared with pure empirical studies conducted without such taxonomy. Hence, the presented taxonomy seems to be a theoretical progress and of practical value in analyzing decision situations, decision-making, and the related research.

REFERENCES

- Gackowski, Z. J. (2004). "Logical interdependence of data/information quality dimensions—A purpose-focused view on IQ," *Proceedings of the Ninth International Conference on Information Quality (ICIQ-04)*, Cambridge, Massachusetts. Retrieved from: <http://www.iqconference.org/Documents/IQConference2004/Papers/LogicalInterdependence.pdf>
- Gackowski, Z. J. (2005b). "Operations quality of information: Teleological operations research-based approach, call for discussion". *Proceedings of the 10th Anniversary International Conference on Information Quality—ICIQ-05* at Massachusetts Institute of Technology (MIT), Cambridge, MA. Retrieved from <http://www.iqconference.org/ICIQ/iqdownload.aspx?ICIQYear=2005&File=OperationsQualityofDataInfo.pdf>
- Gackowski, Z. J., (2006a). "Quality of informing: Credibility - Provisional Model of Functional Dependencies", *Proceedings of the 2006 Informing Science and IT Education Joint Conference*. Salford, UK – June 25-28, retrieved from <http://proceedings.informingscience.org/InSITE2006/ProcGack194.pdf>
- Gackowski, Z. J. (2006b). "Redefining Information Quality and its Measuring: The Operations Management Approach". Accepted for the *Proceedings of the 11th International Conference on Information Quality—ICIQ-06* at Massachusetts Institute of Technology (MIT), Cambridge, MA.
- Liu, L., Chi, L. N., (2002). "Evolutional Data Quality: A Theory-Specific View," *Proceedings of the Seventh International Conference on Information Quality (ICIQ-02)*, Cambridge, Massachusetts.
- March, J. G., "A Primer on Decision Making – How Decisions Happen". Free Press, New York, NY, 1994.
- Oliviera, P., Rodrigues, F., Henriques P., (2005). "A formal definition of Data quality problems," *Proceedings of the 10th Anniversary International Conference on Information Quality—ICIQ-05* at Massachusetts Institute of Technology (MIT), Cambridge, MA. Retrieved from <http://www.iqconference.org/ICIQ/iqdownload.aspx?ICIQYear=2005&File=AFormalDefinitionofDQProblems.pdf>
- Liu, L., Chi, L. N., "Evolutional Data Quality: A Theory-Specific View," *Proceedings of the Seventh International Conference on Information Quality (ICIQ-02)*, Cambridge, Massachusetts.
- Wand, Y., and Wang, R. Y., (1996). "Anchoring Data Quality Dimensions in Ontological Foundations," *Communications of the ACM*, 39(11), 1996, pp. 86-95.
- Wang, R. Y., and Strong, D. M., (1996). "Beyond Accuracy: What Data Quality Means to Data Consumers", *Journal of Management Information Systems (JMIS)*, 12(4), 1996, pp. 5-34.

0 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/proceeding-paper/quality-decision-variables/33053

Related Content

FLANN + BHO: A Novel Approach for Handling Nonlinearity in System Identification

Bighnaraj Naik, Janmenjoy Nayak and H.S. Behera (2018). *International Journal of Rough Sets and Data Analysis* (pp. 13-33).

www.irma-international.org/article/flann--bho/190888

Automated System for Monitoring and Diagnostics Pilot's Emotional State in Flight

Tetiana Shmelova, Yuliya Sikirda and Arnold Sterenharz (2021). *International Journal of Information Technologies and Systems Approach* (pp. 1-16).

www.irma-international.org/article/automated-system-for-monitoring-and-diagnostics-pilots-emotional-state-in-flight/272756

Software Development Life Cycles and Methodologies: Fixing the Old and Adopting the New

Sue Conger (2011). *International Journal of Information Technologies and Systems Approach* (pp. 1-22).

www.irma-international.org/article/software-development-life-cycles-methodologies/51365

Early Warning Model of College Students' Psychological Crises Based on Big Data Mining and SEM

Rui Liu (2023). *International Journal of Information Technologies and Systems Approach* (pp. 1-17).

www.irma-international.org/article/early-warning-model-of-college-students-psychological-crises-based-on-big-data-mining-and-sem/316164

Architecture of an Open-Source Real-Time Distributed Cyber Physical System

Stefano Scanzio (2018). *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 1227-1237).

www.irma-international.org/chapter/architecture-of-an-open-source-real-time-distributed-cyber-physical-system/183836