DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

ONDERZOEKSRAPPORT NR 9612

Representing, Planning and Evaluating Audit Evidence with Belief Function Theory

by

Carine VAN DEN ACKER

Jan VANTHIENEN

Katholieke Universiteit Leuven

Naamsestraat 69, B-3000 Leuven

ONDERZOEKSRAPPORT NR 9612

Representing, Planning and Evaluating Audit Evidence with Belief Function Theory

by

Carine VAN DEN ACKER
Jan VANTHIENEN

D/1996/2376/12

Representing, Planning and Evaluating Audit Evidence with Belief Function Theory

Carine Van den Acker, Jan Vanthienen

Abstract

This paper presents a framework for the representation, planning and evaluation of uncertain audit evidence using a belief-function approach. We only consider the case of a financial statement audit. The framework allows for the modeling of audit risk and audit evidence at different levels of detail, thereby incorporating dependencies at all levels. It provides the information needed to plan an efficient audit and provides means for intermediate and final evaluation. The paper includes a basic introduction to belief-function theory. Then the framework is described in detail, first for aggregation purposes, then for dissaggregation. Next, it is explained how different types of audit evidence should be represented. Finally, we use the audit of the sales and collection cycle and one other fictitious cycle to illustrate representation as well as planning and intermediate and final evaluation. We conclude with a discussion.

Background and Related Work

Dempster-Shafer theory or the theory of belief functions was developed in the late sixties and the seventies. It was called a mathematical theory of evidence, aimed at representing and combining different items of uncertain evidence. Its mathematics are described in Shafer's monograph (1976) that is still a standard reference. More than a decade later, the focus on artificial intelligence (AI) and knowledge based systems redirected the attention to the belief-function theory as a formalism for modeling uncertainty reasoning. This interest from the rapidly growing field of AI inspired many researchers to develop the theory further. Most of the advances stem from the field of AI or are at least aimed at applications in this domain.

As auditing is defined as "the accumulation and evaluation of evidence ..." (Arens and Loebbecke, 1994, 1), this is precisely the type of problem for which the belief-function theory was originally developed. Its potential for auditing was brought to the attention by Shafer and Srivastava (1990). They found that the theory is very well suited to model the auditor's thought process and the true meaning of audit risk, and to take into account the complexity of the structure of audit evidence (Srivastava and Shafer, 1992). Especially the possibility to represent audit evidence and to control audit risk at different levels of detail, is an advantage over other alternatives, like the Audit Risk Model (ARM).

In our opinion, the Belief-function theory and the progress made by AI-researchers entail many promises for the auditing profession. Firstly, because audit plans are no longer considered as static requirements but should allow for feedback from test results, there is a need for more comprehensive methods that can represent and combine these results. Modern computer assisted audit tools (CAATs) are expected to reflect this evolution and to support interactive auditing in which opinions and plans can be revised in the light of new evidence (Hart, 1995). Belief-function theory as it is known today, is well suited to provide a model for the auditing thought process and to allow for computerized support. Secondly, the audit profession can benefit from a comprehensive belief-function framework to integrate evidence and formulate opinions, because it can improve consistency among auditors by reducing judgmental biases (McMillan and White, 1993).

Substantial work has been done by Srivastava e.a.[1] The major limitations of his framework for representing and reasoning with audit evidence were indicated by himself and were taken as a starting point for the framework presented in this paper. We will discuss these limitations and our solution in the last section of this paper.

The purpose of this paper is to present a framework that incorporates the advantages of belief-function theory and does not suffer from the limitations mentioned above. The next section provides an elementary introduction to the theory. In the body of the paper, the framework itself is explained, followed by an illustration. The last section is a discussion.

Basics of Belief-function Theory

Belief functions defined on one frame of discernment

Static aspects

Belief-function theory allows for the expression of belief over a set of possible hypotheses or answers to a question of interest. The set of answers is called *a frame* of discernment (Θ) and includes all possible answers of which only one corresponds to the truth. *Belief* represents the degree of support in favor of one or more hypotheses of the frame. It results from supportive evidence and can be expressed for any subset of the frame. The quantification of belief for any subset of the frame, constitutes a belief function over that frame. Belief is derived from what is called a basic belief assignment (bba) or mass assignment (m). Mass committed to a particular subset is the amount of support in favor of that subset that can not be committed to any smaller set. The belief in a subset A is then calculated as the sum of mass committed to A itself and to all its subsets B, or more formally:

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(B) \ , A \subseteq \Theta$$ **Definition 1**

The sum of all mass committed must be equal to one. In addition to measures of belief, a belief function embeds another interesting measure, namely the plausibility of a subset of hypotheses. The plausibility (Pl) of a subset of hypotheses equals one minus the belief in contradicting hypotheses. Hence, it represents the lack of contradicting evidence. Whereas Bel(A) is the extent to which one beliefs that A *is* true, Pl(A) represents the extent to which one beliefs that A *can* be true or the extent to which the alternative “not A” has not been proven to be true :

$$\mathrm{Pl}(A) = 1 - \mathrm{Bel}(\text{not } A)$$ **Definition 2**

1 See reference list at the end of the paper.

Dynamic aspects

When distinct items of evidence are represented as belief functions on the same frame Θ, they can be combined two by two into one single belief function using *Dempster's rule of combination*(⊕). This rule computes a combined belief function Bel out of two belief functions Bel_1 and Bel_2 by committing the product of the masses to the intersection of the associated subsets. When the intersection is the empty set, there must be conflicting evidence. Combining evidence is aimed at obtaining stronger belief for smaller subsets of hypotheses, ultimately pointing to the single true hypothesis.

Extension : Belief functions defined on a structure of frames

Static aspects

Complex problems do not always allow for representing all hypotheses of interest in one single frame in answer to one single question. Instead, there are interesting questions at different levels of detail with a set of possible answers at each level. Therefore, several frames are defined and related to one another through *refinings*. A refining ω from a frame Θ to a more detailed frame Ω describes for each hypothesis or element of Θ which are the associated hypotheses at the more detailed level, Ω[2]. The refinement for a subset of Θ is then the union of the refinements of its elements. The inverse relation is called a *coarsening*.

Sometimes it is not possible to find a subset of Ω for every element of Θ. Instead, it may happen that only a subset of Θ as a whole can be refined. This is the case of an *incomplete refining*. The purpose of defining several frames and their relationships, is to allow for evidence to be represented at the level of detail that is most appropriate.

Dynamic aspects

Because Dempster's rule of combination can only be used for evidence bearing on the same frame, and because the structure of frames results in evidence represented as belief functions on different related frames, belief must be propagated through the structure in order to combine the evidence. For that purpose, transfer mechanisms are based on refining and incomplete refining relationships and describe how belief committed to hypotheses at a particular level of detail can be transferred to associated hypotheses at levels of more or less detail[3] . Hence, evidence represented on the same frame can be combined using Dempster's rule and evidence represented on different frames can be transferred to the same level and then be combined using Dempster's rule.

---

2 The refinements of different hypotheses may not be empty and may not overlap (their intersection must be empty) and their union must equal Ω. In short, a refining relation from Θ to Ω defines a *partition* on Ω.

3 Explaining the principles of transfer is beyond the scope of this paper. In short, downward transfer is done by assigning mass committed to a subset of the coarsening to its refinement. Upward mass transfer is done by assigning mass committed to a subset of the refinement to the smallest subset of the coarsening of which the refinement comprises the subset to transfer from.

Structure of Frames for the Financial Statement Audit

Evidence Aggregation

In case of a financial statements (FSs) audit, the hypotheses of interest are concerned with the fairness of the FSs. As the auditor will ultimately communicate his opinion through a standardized report, his purpose is to find out the true state of the financial statements and consequently the appropriate report to issue. The answer to the question of fairness of the financial statements is one of the following :

- the FSs are materially misstated (MM_{FS})
- the FSs are immaterially misstated (IM_{FS})
- the FSs are correct (C_{FS})

When MM_{FS} is found to correspond to the truth, an adverse opinion or report will need to be issued. In case C_{FS} is judged to be the true hypothesis, the opinion will be unqualified. If IM_{FS} is believed to be the truth, the opinion may be qualified using different wording or an additional paragraph to draw the attention to the misstatement that was judged to be immaterial to the FSs as a whole, but still too large to be labeled correct. In case IM_{FS} is the true hypothesis, the auditor may of course still want to decide to issue an unqualified opinion depending on the particular circumstances and the type of immaterial error. This set of possible answers or hypotheses of ultimate interest, is the top level frame (level 1, L1), named FS, in **Figure 1** below.

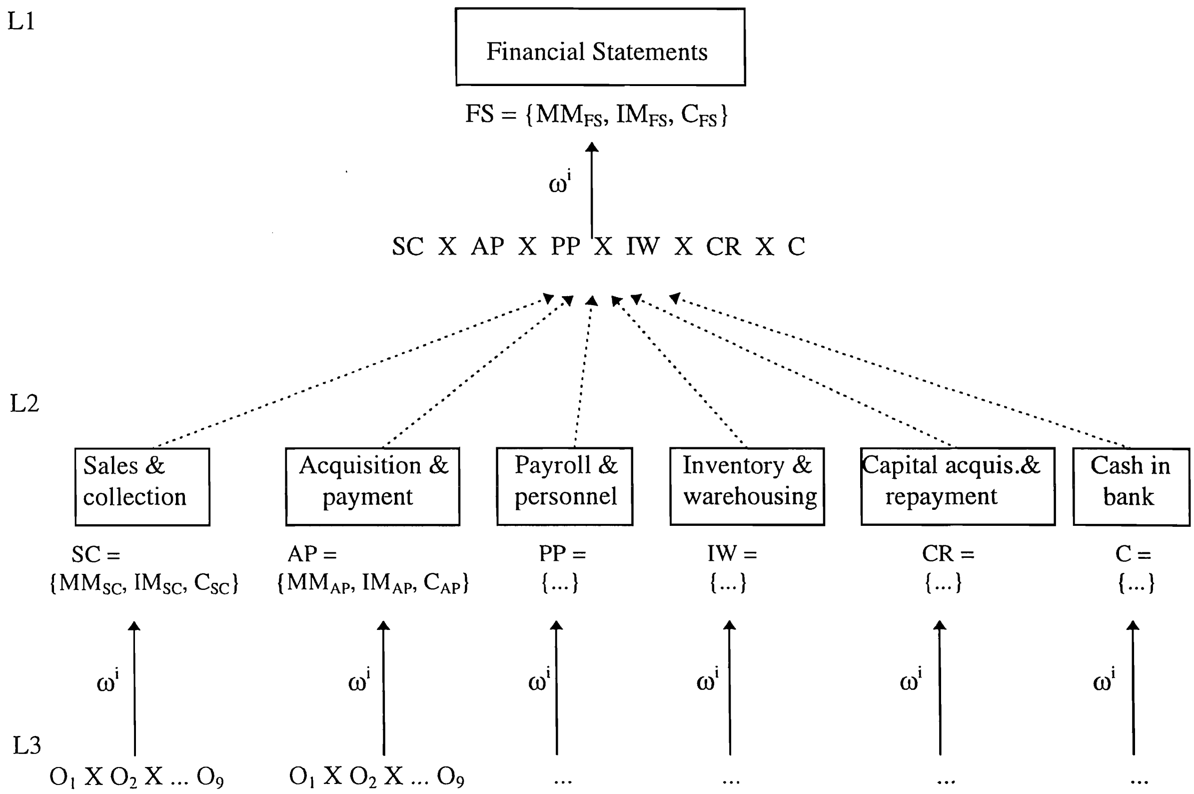

Figure 1 : Structure of Frames for the Financial Statement Audit for Evidence Aggregation

The overall question of fairness of the FSs is answered by segmenting the audit in cycles and judging the fairness of each cycle. Hence, there is a question of fairness for each individual cycle and an associated frame of possible answers. These cycles and their corresponding frames are represented in **Figure 1** at L2. For example, the question of fairness of the sales and collection cycle, has three possible answers :

- the sales and collection cycle is materially misstated (MM_{SC})
- the sales and collection cycle is immaterially misstated (IM_{SC})
- the sales and collection cycle is correct (C_{SC})

MM_{SC} corresponds to the situation in which test results lead to the conclusion that errors are larger than tolerable error ; IM_{SC} represents the conclusion of the presence of errors that are not material by themselves, but can become material when combined with immaterial errors of other cycles ; C_{SC} corresponds to the conclusion that errors will not exceed minimal error. The set of these three hypotheses, SC, is the frame of interest associated with the sales and collection cycle. In a similar way, AP, PP, IW, CR and C are defined at L2.

The collection and evaluation of evidence per cycle will result in a combined belief function on the frame of that cycle. In order to obtain an opinion on the FSs as a whole, the mass supporting a particular subset of hypotheses at the cycle level must be combined with results from other cycles and then transferred to subsets of hypotheses at the FSs level. For this purpose, the product set of the cycle frames is used as a common frame : SC X AP X PP X IW X CR X C (represented between L1 and L2 in **Figure 1**).

The product set contains all possible combinations of hypotheses for each cycle, e.g. (C_{SC}, IM_{AP}, C_{PP}, C_{IW}, C_{CR}, IM_C) is one of its elements meaning : "sales and collection is correct, and acquisition and payment is immaterially misstated, and ..., and cash in bank is immaterially misstated". In our case, the product set of the six cycle frames contains 3^6 or 729 elements. Fortunately, the basic belief assignment (bba) defined on the product set will not be so complex in practice[4].

In order to express a belief function, Bel, defined on a frame at the cycle level as a belief function, Bel^*, defined on the product set, the *cylindrical extension* of all elements with positive mass (focal elements) of Bel must be found, i.e. [5] :

The cylindrical extension of a focal element A in the frame of discernment Θ to a focal element A^* in $\Theta X \Omega$ is defined as : $A^* = \{(x, y) | x \in A, y \in \Omega\}$ **Definition 3**

To illustrate, assume that the result of auditing the sales and collection cycle results in a simple support function with focal element $\{IM_{SC}\}$. The cylindrical extension of this focal element is then $\{(IM_{SC}, y, z, u, v, w) | y \in AP, z \in PP, u \in IW, v \in CR, w \in C\}$ which is the focal element of the belief function defined on the product set. When belief functions defined at the cycle level all have been extended to belief functions defined on the product set, they can be combined into one belief function defined on that product set. The cylindrical extensions are shown in **Figure 1** by the dotted arrows.

The FSs as a whole are considered to be correct if all cycles were found to be correct. Hence, the support for hypothesis $\{C_{FS}\}$ is implied by the support of only one element of the product

4 See illustration at the end of the paper.
5 Hau, H.Y. and Kashyap, R.L. (1990), p. 49.

set, namely[6] : (C_{SC} and C_{AP} and C_{PP} and C_{IW} and C_{CR} and C_C). However, the material misstatement of FSs is concluded whenever there is at least one materially misstated cycle, assuming that the material misstatement of any complete cycle causes material misstatement of the FSs as a whole. Hence, all product set elements containing at least one co-ordinate with value MM correspond to $\{MM_{FS}\}$. All remaining product set elements have at least one co-ordinate with IM value and no co-ordinate with MM value. They do not correspond to the FSs' hypothesis "IM_{FS}", but to the subset of hypotheses $\{MM_{FS}, IM_{FS}\}$ because immaterial misstatements in several cycles can cause FSs to be immaterially misstated, but can also lead to the conclusion of material misstatement at the FSs level when the combined misstatement of all cycles exceeds materiality[7]. This relation between FS and the product set of cycles, is an incomplete refining and is represented in **Figure 1** by the full arrow. This kind of refining is only to be used for transferring mass (or belief) upwards, i.e. for evidence aggregation.

The third level in **Figure 1** is the level of audit objectives, representing the idea that the question of fairness of the cycle is evaluated using subquestions concerned with the achievement of the nine balance-related audit objectives (See Arens and Loebbecke, 1994, p.146). The frame of interest at the level of audit objectives is the product set of the frames O_i, each O_i representing the possible answers to the question of achievement of objective i. For example, $O_3 = \{MM_{O3}, IM_{O3}, C_{O3}\}$ means that with respect to the third objective a particular cycle can either be materially misstated, immaterially misstated or can be correct[8]. In contrast to L2 where a separate frame was used for each cycle to discern the evidence, the frame of interest at L3 is the product set of individual objective frames instead of separate frames for each objective. The need for using the product set to discern all the evidence collected with respect to audit objectives, comes from the fact that evidence frequently bears on several objectives at the same time. Hence, the belief function induced by the evidence should be defined on a frame that incorporates all objectives. The product set discerns all evidence that bears on whatever subset of objectives within a cycle, but for nine audit objectives it contains 3^9 or 19683 elements which gives too large a complexity. Hence, some of its flexibility in representing any kind of evidence shall be sacrificed in order to keep complexity reasonable. This will be accomplished by splitting the set of objectives in majors and minors and referring to minors as one single objective, named "others". Assuming k major objectives and n-k minor objectives, the product set contains 3^{k+1} elements. This approach agrees with common practice.

Similar to the product set of cycle frames, the product set of audit objective frames can be partitioned into subsets of elements that lead to the same hypothesis at the level above. Consequently, we can define an elementary refining for each cycle that relates each partition of the frame at L3 to the appropriate hypotheses at L2. Though the actual refining will be dependent on the cycle it is concerned with, all possible refinings will fall into the pattern of the approach chosen in this text, being :

- all elements of the objectives product set with an MM value for at least one of the k major objectives, constitute the first partition that is associated with $\{MM_{cycle}\}$;

---

[6] Co-ordinates of elements of the product set are separated by "and" instead of "," to avoid confusion with subsets of hypotheses.

[7] Reviewing immaterial errors on the unadjusted error worksheet and deciding whether or not they exceed materiality, will provide evidence that is to be included in a next stage.

[8] Symbols at L3 will not be indexed for the cycle they refer to, because it would complicate notation whereas the cycle will be clear from the context when needed.

- all elements of the objectives product set with a C value for the k major objectives and a C or IM value for "others", constitute the second partition that is associated with $\{C_{cycle}\}$;
- all remaining elements constitute the third partition that is associated with $\{IM_{cycle}, MM_{cycle}\}$ because of the fact that the composite of immaterial misstatements can either become material or remain immaterial at a higher level.

Again, these incomplete refinings are unidirectional, as indicated by the full arrows between L2 and L1 in **Figure 1**. They should only be used for upward mass transfer.

A similar way of reasoning must be defined in order to decide for what combinations of values for the minor objectives, "others" obtains the values C, IM or MM. The way of reasoning must be adapted to the characteristics of the cycle and the preferences of the auditor. To illustrate, when no distinction is made within the subset of minor objectives, the following reasoning could be applied :

- When there is support that all minor objectives are correct, then $\{C_{other}\}$ is supported too;
- when there is support that no minor objective is materially misstated, and at least one of them is immaterially misstated, then $\{IM_{other}\}$ is supported too;
- when there is support that at least one minor objective is materially misstated, then $\{MM_{other}\}$ is supported too.

However, audit evidence usually bears on audit objectives directly related to balances and transactions and is only related to cycles in an indirect way. Furthermore, substantive tests usually concentrate on balance sheet accounts with the fairness of income statement accounts obtained as a by-product. In addition, accounts and transactions of minor importance can often remain untested, except for analytical procedures. Finally, it is common practice to perform extensive testing on the "central" account of the cycle, i.e. the account that is involved in almost all transactions, because fairness of that account highly supports fairness of minor accounts because of the double-entry bookkeeping system. The major streams of transactions influencing the balance of the central account are also audited extensively to support the fairness of that account, whereas less important transactions can remain untested. All these reasons allow the question of fairness of a cycle to be simplified to the question of fairness of its central account. Therefore, the audit objectives at L3 in **Figure 1** are in fact the balance-related audit objectives of the central account.

2 *Evidence Disaggregation*

The audit process also entails the collection of more general or higher level evidence that will influence detailed testing. More precisely, assessments of inherent risk and control risk can bear on a cycle as a whole, instead of on audit objectives. In that case, evidence induces belief on higher level frames, that should be transferred to "equivalent" hypotheses at the lowest level to be combined with detailed testing evidence. This problem does not exist when all evidence is directly expressed in terms of major accounts and audit objectives, thereby inducing belief functions defined on the objectives product frames. When this is the case, the issue of a downward mass transfer becomes irrelevant and the framework of **Figure 1** can handle the evidence representation and propagation for the entire audit process. When it is not, the relations in the framework of **Figure 2** are needed for dissaggregating audit evidence from higher to lower levels.

When evidence is gathered and combined at the FSs level (L1), it induces a belief function defined on FS, i.e. it induces support for one or more hypotheses of the top level frame. For each singleton hypothesis in FS, one can define which cycle level hypotheses should be affected by its support, more precisely :

- if there is support that $\{MM_{FS}\}$ contains the true state, then there is also support that at least one cycle is materially misstated;
- if there is support that $\{C_{FS}\}$ contains the true state, then there is also support that each cycle must be correct;
- if there is support that $\{IM_{FS}\}$ contains the true state, then there is also support that no cycle is materially misstated and at least one cycle is immaterially misstated.

Hence, each element of the frame FS can be associated with a subset of elements of the product set of cycle frames. The association defined by the three rules mentioned, is a refining (indicated by the full arrow in **Figure 2** and will be used to transfer mass from L1 down to the product set.

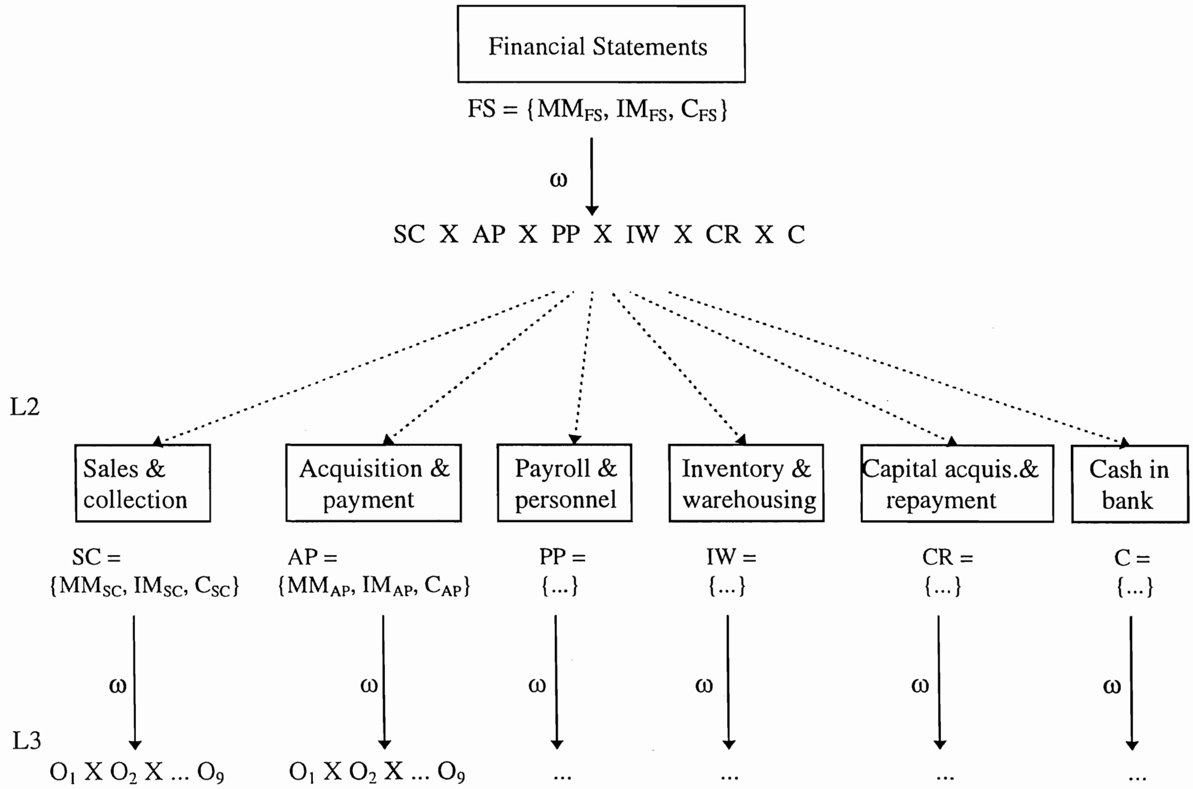

Figure 2 : Structure of Frames for the Financial Statement Audit for Evidence Disaggregation

When evidence is represented using a belief function defined on the product set of cycle frames, it cannot yet be combined with evidence that is collected per cycle. Therefore, support expressed in terms of combinations of cycles (product set), must be transformed into support that bears on the individual frame of each cycle. In order to express a belief function, Bel, defined on the product set of cycle frames, as a belief function Bel_*, defined on one particular cycle frame, the *projection* of all focal elements must be found, i.e.[9] :

The projection of a focal element B in $\Theta X \Omega$ to the frame of discernment Θ, is a focal element B_* where : $B_* = \Pi_\Theta(B) = \{x \mid (x,y) \in B\}$ **Definition 4**

[9] Hau, H.Y. and Kashyap, R.L. (1990), p. 49.

To illustrate, assume that inherent risk assessment induces a belief function with focal element $\{C_{FS}\}$ which is refined to $\{C_{SC}, C_{AP}, C_{PP}, C_{IW}, C_{CR}, C_C\}$. Projecting this focal element on the frame of the sales and collection cycle, SC, induces a belief function on SC with focal element $\{C_{SC}\}$. Projections in **Figure 2** are indicated by the dotted arrows.

After mass is transferred from L1 to L2, it can be combined with evidence that bears directly on hypotheses at L2 (e.g. control risk assessments, evidence from tests of controls), inducing a single belief function on each cycle frame. Knowing which hypothesis is supported by evidence for the cycle as a whole, leads to knowing which hypotheses at the objective level should also be supported. More precisely :

- if there is support that $\{MM_{cycle}\}$contains the true state, then there is also support that there is a material misstatement with respect to at least one of the k major objectives;
- if there is support that $\{C_{cycle}\}$ contains the true state, then there is also support that the k major objectives are evaluated to be correct and that the group of minor objectives is evaluated as immaterially misstated or correct;
- if there is support that $\{IM_{cycle}\}$ contains the true state, then there is also support that
  * none of the k major objectives are evaluated to be materially misstated and at least one of them is evaluated as immaterially misstated, or
  * all k major objectives are evaluated to be correct and the group of minor objectives is evaluated to be materially misstated.

Again, the association of each of the elements of a coarser frame (L2) with a partition of elements of a finer frame (L3) is a refining, which will be used to transfer mass downwards to L3. At L3, detailed evidence will be collected and represented using the product set of objectives frames. It will be combined at that level. When sufficient evidence is collected, it will be aggregated using the framework of **Figure 1**.

Evidence Representation

The framework presented in the previous sections allows for the representation, aggregation and disaggregation of audit evidence. This section will explain how the different types of audit evidence can be represented in the framework[10].

Evidence outcomes can be represented in more than one way, each of which can be defended. The point of view chosen in this text, is the common understanding that the auditor's ultimate concern is the plausibility of material misstatement after the audit is completed. All evidence is collected with the purpose of reducing this plausibility, and hence represented in most cases as support for the complement of material misstatement. This way of representing evidence, together with the fact that material misstatements are mostly corrected, makes that the normal scenario is the combination of corroborating pieces of evidence narrowing the focal elements of the combined belief function.

[10] In this paper, one item of evidence is assumed to provide the same degree of support to all variables of the frame of discernment. Recently, Srivastava and Cogger (1995) developed a heuristic algorithm for representing evidence that gives different support to different variables. For numerical and notational simplicity, this complication for the bba is not considered in this paper, although the framework can handle this case as as well.

Inherent Risk Assessments

By examination of the factors that determine inherent risk (IR), the auditor will quantify IR, i.e. the plausibility of material misstatements due to inherent risk factors. Using **Definition 1 and 2** $Pl(\{MM\}) = IR$ is represented as :

$Bel(\{IM, C\}) = 1\text{-}IR$
$m(\{IM, C\}) = 1\text{-}IR$

There is no evidence to commit support to {IM} or {C} separately that is not committed to {IM, C}, because only the presence or absence of material misstatements is focused at this stage. Remaining mass (equal to IR) is committed to the frame.

Cycle Level

In case the auditor can only specify IR per cycle, it will be represented at L2 as :

$m(\{IM_{cycle}, C_{cycle}\}) = 1\text{-}IR_{cycle}$

To allow IR to influence detailed testing, it must be propagated from L2 to L3. To illustrate, let us assume a sales and collection cycle with central account "accounts receivable (A/R)" and two major audit objectives "existence" and "realizable value" (remaining objectives are labeled "others"). Further assume that inherent risk for sales and collection is set at 70 %, then : $m(\{IM_{SC}, C_{SC}\}) = 0.3$

Mass is propagated to the product set of objectives (L3), in this case $O = O_{exist} X O_{realiz} X O_{other}$ using the refining relation[11] :

$m(\{(x, y, z)|x, y\in\{IM,C\}) = 0.3$

Projecting this belief function on individual objectives frames, would induce belief functions with respective focal elements $\{MM_{others}, IM_{others}, C_{others}\}$ and $\{IM_i, C_i\}$ for major objective i. Hence, $Pl(MM_i) = 0.7$ for major objective i, but $Pl(MM_{others}) = 1$ (and not 0.7), caused by the definition of the incomplete refining.

Audit Objectives Level

It is also possible that IR is specified at the level of major and minor objectives related to the central accounts. Continuing the example, assume that the auditor has evidence from source 1 that plausibility of material error with respect to existence and minor objectives is 70% and evidence from source 2 that plausibility of material error with respect to existence only is 60%. This kind of evidence is treated as distinct evidence, because of the simplifying assumption that evidence with a different degree of support does not come from the same source. This evidence is discerned by the product set $O = O_{exist} X O_{realiz} X O_{other}$:

source 1 : $m_1(\{(x, y, z)|\ x,z\in\{IM,C\}) = 0.3$
source 2 : $m_2(\{(x, y, z)|\ x \in\{IM,C\}) = 0.4$

Combining with Dempster's rule yields $m = m_1 \oplus m_2$:

$m(\{(x, y, z)|\ x,z\in\{IM,C\}) = 0.3 \times 0.4 + 0.3 \times 0.6 = 0.30$
$m(\{(x, y, z)|\ x\in\{IM,C\} = 0.7 \times 0.4 = 0.28$

Projecting on individual objectives frames, yields $Pl(MM_{exist})=0.42$, $Pl(MM_{realiz})=1$, $Pl(MM_{other})=0.7$, representing respectively the corroboration of two items of evidence bearing

[11] The first co-ordinate refers to "existence", the second to "realizable value", the third to "others".

on existence, the fact that there was no evidence concerning realizable value, and the presence of one item of evidence bearing on minor objectives.

Control Risk Assessments

Assessing control risk involves the quantification of the plausibility that material misstatements in FS balances remain undetected because of the failure of client's controls on transactions. Auditing Standards prescribe that control risk may be set at a value lower than one, only when tests of controls are performed and support this value. Test of controls are performed per cycle on the major transactions that are related to the central account of that cycle, and provide evidence with respect to one or more transaction-related audit objectives. Evidence can be completed by a quantification of control risk for "other transactions" in the cycle. Because the framework only includes balance-related audit objectives, conclusions reached must be reformulated in terms of balance-related objectives. This switch is easy because each transaction-related objective has an equivalent balance-related objective to which it is either directly or inversely related, depending on the decreasing or increasing effect of the transaction on the account's balance. To illustrate, A/R are affected by two major transactions, being sales that increase A/R and cash receipts that decrease A/R. The switch from transaction-related objectives to balance-related objectives is done by :

Sales		**A/R**		**Cash receipts**
existence	→	existence	↙	existence
completeness	→	completeness	↖	completeness
accuracy	→	accuracy	←	accuracy
classification	→	classification	←	classification
timing	→	cutoff	←	timing
posting& summarization	→	detail tie-in	←	posting& summarization

Table 1 : Relation of Transaction-related Objectives (Sales and Cash Receipts) to Balance-related Objectives (A/R)

If the initial control risk assessment could be adopted without performing supportive tests, it could be represented similar to inherent risk as a belief function with focal element {IM, C} and associated mass equal to one minus assessed CR. In fact, collecting evidence to support the value of CR is a peculiar matter because it means that evidence must be found to support the amount of belief that is given to {IM, C}, not to support the hypotheses of ultimate interest themselves. Tests of controls will therefore render *support for the amount of support* that can be committed to {IM, C} or stated differently, tests will be used to conclude whether the assessed value of CR is the appropriate amount of belief to be given to {IM, C}.

Representing evidence from tests of controls requires a different approach when samples are used than when tests are performed on a non-sampling basis.

Non-sampling Evidence

"Inquiry of personnel" and "observation" are two types of tests of controls that are not done on a sample basis. They lead to conclude whether the assessed level of CR, say e.g. 0.8, is justified or not. If it is justified, 0.8 can be accepted as CR and represented by committing 0.2 to {IM, C}. Evidence rarely supports hypotheses of interest with certainty because test procedures are seldom perfectly reliable (say reliability is 0.9). Hence, it is preferable to discount the accepted support of 0.2 for this risk, by multiplying by 0.9 or in general, a discounting factor d :

$m(\{IM, C\}) = 0.2 \times d = 0.18$

thereby increasing the control risk to :

$Pl(\{MM\}) = 1 - 0.2 \times d = 0.82$

With or without discounting, inquiries and observations can be used to back the assessed quantification of the strength of controls with respect to whatever mix of objectives. Different quantifications are considered to come from distinct items of evidence, each one inducing a belief function on the product set of k major and n-k grouped minor objectives, in a similar way as was explained for inherent risk. For example, evidence bearing on CR for completeness and timing of cash receipts, bears on existence (major objective) and cutoff ("others") of A/R and can be represented as :

$Pl(\{(x, y, z)| x,z\in \{MM\}) = 1 - (1 - CR) \times d$ or

$m(\{(x, y, z)| x,z\in \{IM,C\}) = (1 - CR) \times d$

Although these procedures are subjective, the auditor will still need to formulate a decision rule in advance, i.e. he has to decide how "bad" the outcomes of inquiries and observations may be and still support the 0.8 assessment of CR. Findings must then be evaluated using this rule. When outcomes are more positive, the auditor can maintain the 0.8 assessment; when they are worse, he must increase CR to represent the outcomes. The result can then be discounted to reflect the less than perfect assurance given by the test procedure.

Sampling Evidence

"Examination of documentation" and "reperformance" are two types of tests of controls that are performed on a sample basis in order to find support for the assessed CR. The sampling method used is attributes sampling, yielding the sample exception rate, r (rate of control failures) that can be used to calculate the assurance associated with intervals comprising the population error rate, R. When tests are performed on a sample basis, the need to account (discount) for the risk of incorrect acceptance is more apparent than in the non-sampling case, although also in this case, discounting is seldom encountered. The result of tests of controls using samples, is the level of CR that will be adopted. The degree of assurance associated with that conclusion must be sufficient (the common 95% assurance or α=5%), but its precise value is usually ignored.

Again the auditor must formulate a decision rule in advance : he must set the tolerable exception rate for the population (TER) that corresponds to his CR assessment, say 0.8.. The auditor must set the level of control failure rate on transactions that would cause 80% CR. When TER is set at e.g. 6 %, it means that controls that fail for 6 % or less of transactions, would make the auditor accept control risk of 80 %.

After testing the sample, r is known, on the basis of which P(R<TER) can be calculated. When the result is at least required assurance (1-α) or 0.95, the 0.8 level of CR will be adopted. Else, sample size must be increased and assurance recomputed, or CR must be

reassessed and retested. In any case, when the assurance with which the final value of CR is concluded is at least 95 %, CR is kept as the support for {IM, C} and will influence further testing. The information that this conclusion was reached with an assurance of at least 95 % is often ignored in further computations. As the degree of assurance of the sampling evidence is valuable information, it would be better to discount the resulting CR value with the factor d equal to P(R <TER) to represent the sampling risk:

$Pl(\{MM\}) = 1 - (1 - CR) \times d$ or
$m(\{IM,C\}) = (1 - CR) \times d$

Substantive Evidence

Substantive Tests of Transactions

In case of tests of controls, the auditor examines the failure rate of particular client's controls. When controls fail, they only indicate a possibility of misstated transactions because transactions can still be stated correctly. In case of substantive tests of transactions, the auditor examines the amount of misstatements in transactions. Deviations found definitely are misstatements and not just an indication of possibility.

Similar to tests of controls, attributes sampling is the sampling method that is often used for substantive tests of transactions. The auditor must define in advance which amount of misstatement in the sample will be considered as an exception and which population rate of these exceptions can be tolerated (TER) and still makes the conclusion of "no material misstatements" acceptable. Through testing, the auditor will obtain r which allows him to compute P(R<TER) or the assurance that is associated with the "no material misstatements" conclusion. If the assurance is below the level of required assurance, sample size can be increased or possibly TER can be revised. With or without corrections or adjustments made, final result of testing will be that P(R<TER) is greater than P(R>TER). The evidence is then represented in the following bba (for a formal description see Appendix I) :

$m(\{IM,C\}) = P(R < TE)$
then $Pl(\{MM\}) = P(R > TE)$

Because substantive tests of transactions bear on audit objectives related to major transactions, results will be represented in the product set of objectives frames.

Tests of Details of Balances

Tests of details of balances are aimed at establishing the monetary correctness of account balances. We will assume the use of monetary unit sampling (MUS). Unlike attributes sampling, MUS yields the sample amount of misstatement instead of error rates. The auditor must decide in advance which amount of misstatement in the population (x) is so severe that the conclusion of presence of material misstatement should be drawn (i.e. tolerable error,TE), and which amount of misstatement is so small that the conclusion of correctness is appropriate (i.e. minimal error, ME). Misstatements in between these boundaries are considered immaterial by themselves, but can become material when combined with other immaterial errors. TE and ME can be set at different levels for overstatement and understatement errors, then indicated as TE^+, TE^-, ME^+ and ME^-.

Sample data are used to assess the amount of misstatement in the population. With respect to that population amount of misstatement , the auditor is interested in three intervals, being :

(1) $P(ME^- < x < ME^+)$ "correct"
(2) $P(TE^- < x < ME^- \cup ME^+ < x < TE^+)$ "immaterial"
(3) $P(x < TE^- \cup x > TE^+)$ "material"

Depending on the possible ranking of (1), (2) and (3) we can distinguish between three cases.

Case I :
Under normal circumstances, (1) will be the largest, followed by (2) and then by (3). For the evidence to be sufficient, the auditor will require that the assurance that there are no material misstatements ((1)+(2)) is at least the assurance needed. When this is the case, evidence can be represented by the bba :

$m(\{C\}) = P(ME^- < x < ME^+)$
$m(\{C, IM\}) = P(TE^- < x < ME^- \cup ME^+ < x < TE^+)$
$m(\{C, IM, MM\}) = P(x < TE^- \cup x > TE^+)$

In case the assurance required is not provided, the auditor will extend testing, revise threshold values etc. in order to increase assurance and recompute the intervals.

Case II :
In other circumstances, the sample may contain such an amount of misstatements that (3) is ranked second or even first. The auditor will then ask for adjustment bookings or even correction of the population by the client and will recompute the intervals (in case of adjustment) or reperform testing (in case of correction) in order to reconsider his conclusions. When corrective action is not taken, the erroneous situation that was discovered continues to exist and the auditor will have to present this evidence by :

$m(\{MM\}) = P(x < TE^- \cup x > TE^+)$
$m(\{C, IM, MM\}) = 1 - P(x < TE^- \cup x > TE^+)$

Because other types of evidence (inherent risk, tests of controls, substantive tests of transactions) all have been represented as the plausibility of material misstatement or the support for {IM, C}, conflict will arise.

Case III :
Although rather uncommon, there may be circumstances in which (2) is the largest, followed by (1) and then (3). When the assurance provided by (1)+(2) obeys the requirement, then the probability of material misstatement is low enough to consider the evidence as sufficient. In that case, the bba of the consonant belief function induced by the evidence is :

$m(\{C, IM\}) = P(TE^- < x\ TE^+)$
$m(\{C, IM, MM\}) = P(x < TE^- \cup x > TE^+)$

When the auditor foresees that the high occurrence of immaterial misstatements will induce material misstatements when aggregated, he can decide to increase sample size, revise threshold values or take corrective actions such that reconsideration of the evidence can lead to Case I.

Analytical Procedures

Analytical procedures can induce evidence at all levels, because they are evidence for overall reasonableness as well as for final review and can also provide substantive evidence. When analytical procedures are performed using ratio analysis or even simple comparisons, the amount of support that is associated with their conclusions is determined subjectively as in the case of inherent risk. When more sophisticated techniques, e.g. regression analysis, are used, it may be possible to compute the support. This issue is however beyond the scope of this text.

Aggregation of Unadjusted Immaterial Misstatements

During evidence collection, misstatements smaller than tolerable error but larger than minimal error are listed on the unadjusted error sheet. The possibility that misstatements that are immaterial in separation can become material in combination, is modeled in the framework by the incomplete refining that associates {IM} at the lower level with {IM, MM} at the higher level. As soon as evidence collection is considered complete at the lower level and mass is transferred to the higher level, it can be decided whether the aggregate misstatement is material or not.

This decision is not really evidence like the previous four types of audit evidence, as it does not add new support to the frame of interest. The support for {IM, MM} is already known from testing procedures ; it only remains to be decided which of both hypotheses will be the one that is supported by the evidence. Hence, this decision should not be presented as a bba because it does not induce new support. Instead it causes a *shift of mass* from {IM, MM} to either {IM} or {MM}. Details of that shift depend on the belief function that is obtained after combination at the lower level and before propagating upwards.

When all substantive evidence induces belief functions with focal elements {C} and {C, IM}, then focal elements remain the same after combination with inherent risk and control risk assessments. The transfer using the incomplete refining will assign m({C,IM}) "temporarily" to m(C, IM, MM)[12]. After examining the unadjusted error sheet, this mass will be shifted to :

- {C, IM} when the aggregate error is below TE
- {C, IM} when the aggregate error is above TE, but corrective action is taken
- {MM} when the aggregate error is above TE and no corrective action is taken

Substantive evidence combined with inherent risk and control risk assessments can also induce a belief function with other focal elements in which case similar shifts are necessary.

The representation of evidence in the framework as well as the use of the relations, is illustrated in the next section.

[12] In case immaterial misstatements only involve minor objectives, then mass is transferred from {C, IM} to {C}.

Illustration : Audit of the Sales and Collection Cycle

In this section the representation of different types of evidence is illustrated for the sales and collection cycle. The way of proceeding is the same for other cycles. When the audit of this cycle is considered complete, results are combined with those of a fictitious cycle in order to illustrate the reasoning for evidence aggregation at the cycle level. On the basis of these two cycles, conclusions w.r.t. the fairness of the FS as a whole are drawn. To keep the example manageable, the illustration assumes the performance of only a limited number of audit tests.

Description

The auditor decides that A/R is the central account which will be audited intensively, because A/R is involved in nearly all transactions. Tests for other accounts can be limited (e.g. analytical procedures), and concentrate on objectives that are not influenced by the double-entry system (e.g. posting and summarization). The major transactions of the cycle are sales and cash receipts. Other transactions are considered of minor importance (e.g. charge-off of uncollectible accounts, sales returns and allowances) or will be part of another cycle (e.g. cash sales). The audit objectives that are considered of most importance for A/R are decided to be existence, realizable value and cutoff.

Representing Inherent Risk Assessment

After evaluating the nature of the client's business, it was decided that inherent risk is low for all objectives in general. Results from allowances for doubtful accounts of previous audits indicate a medium level of risk for realizable value and provides no evidence for the remaining objectives. Medium and low are quantified as IR equal to 0.8 and 0.6 respectively. The evidence is represented on the product frame of objectives $O_{SC} = O_{exist} \times O_{realiz} \times O_{cutoff} \times O_{other}$:

$m_1(\{(x,y,z,u) \mid x,y,z,u \in \{IM,C\}\}) = 0.4$
$m_2(\{(x,y,z,u) \mid y \in \{IM,C\}\}) = 0.2$

combined evidence yields :

$m(\{(x,y,z,u) \mid y \in \{IM,C\}\}) = 0.12$
$m(\{(x,y,z,u) \mid x,y,z,u \in \{IM,C\}\}) = 0.4$

Projecting on individual objectives frames yields Pl(MM) after considering IR as shown in the upper half of **Table 2**. It shows that in this case of simple support functions, the combination of plausibility due to inherent risk factors equals the product of individual plausibilities when projected on one single frame, which is the same result as when the ARM[13] is applied at one level for one objective.

[13] Recall that ARM prescribes : AAR = DR x CR x IR

Pl(MM)	Existence	Realizable value	Cutoff	Others
IR - business	0.6	0.6	0.6	0.6
IR - previous audit	1	0.8	1	1
IR	**0.6**	**0.48**	**0.6**	**0.6**
CR - sales	0.7	1	1	0.7
CR - cash receipts	0.6	0.6	0.6	0.6
CR	**0.42**	**0.6**	**0.6**	**0.42**
CR⊕IR	***0.252***	***0.288***	***0.36***	***0.252***

Table 2 : Inherent and Control risk for A/R

Representing Control Risk and Tests of Controls

The client's internal control structure is examined and strengths and weaknesses identified. Based on that evaluation, control risk is assessed at 0.7 for sales with respect to all objectives and at 0.6 for cash receipts. Tests of controls for sales are only performed for existence, accuracy and timing. Tests of controls for cash receipts are performed for all objectives[14]. After increasing the size of a few samples, the required 95 % assurance was met. Hence, the assessed CR can be adopted for all objectives in case of cash receipts, but only for the three objectives tested in case of sales[15].

Results are :

$m(\{(x,y,z,u) \mid x,u \in \{IM,C\}\}) = 0.3$ for sales

$m(\{(x,y,z,u) \mid x,y,z,u \in \{IM,C\}\}) = 0.4$ for cash receipts

and after combination :

$m(\{(x,y,z,u) \mid x,u \in \{IM,C\}\}) = 0.18$

$m(\{(x,y,z,u) \mid x,y,z,u \in \{IM,C\}\}) = 0.4$

At this point the combined support from inherent and control risk factors can be determined by combining both bba's on the product set, yielding :

$m(\{(x,y,z,u) \mid x,y,z,u \in \{IM,C\}\}) = 0.64$ $m(\{(x,y,z,u) \mid x,u \in \{IM,C\}\}) = 0.0864$

$m(\{(x,y,z,u) \mid x,y,u \in \{IM,C\}\}) = 0.0216$ $m(\{(x,y,z,u) \mid y \in \{IM,C\}\}) = 0.0504$

Again, this bba can be projected on the individual objectives frames. Results are shown in the lower half of **Table 2**. They indicate once more that combined plausibility expressed with respect to one single objective, equals the product of IR and CR as formulated in the ARM.

Even though projection on individual frames can provide a check on the results, these individual plausibilities are less valuable than the information contained in the bba expressed on the product set. It is indeed more important to know how plausible misstatements are with respect to a number of objectives at the same time instead of for each objective separately.

---

14 For examples of tests of controls for sales transactions and cash receipts, see Arens and Loebbecke, p. 354 and 360.

15 For numerical simplicity, the accepted level of CR is not discounted for the assurance of testing.

From the belief function, one can compute that the plausibility of material misstatements with respect to e.g. :

existence, realizable value and others equals 0.3384,
existence, cutoff and others equals 0.36,
existence and others equals 0.252.

This kind of information is not contained in belief functions defined on individual objectives frames and is valuable because detailed testing often provides evidence with respect to several objectives at the same time. When a substantive test provides evidence on existence and realizable value at the same time, the sample size can be determined on the basis of the combined assurance that is needed. For example, when 0.05 is acceptable audit risk and 0.3384 remains before detailed testing, then substantive testing should provide a 85.2% assurance[16] that no material misstatements are present. In case only the information in **Table 2** was available, then two sample sizes would be computed, one for existence and a different one for realizable value, whereas both objectives can be checked at the same time. When there is no information on the "combined" level of assurance required, then there is likely to be over- or under audit.

Representing Substantive Evidence

Based on the results of evidence already collected at this stage, the auditor decides to perform 7 types of substantive tests for A/R[17], each one covering one or more audit objectives as shown in **Table 3**. For a description of these tests, see Appendix II.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Existence		x	x			x	
Realizable value		x			x		
Cutoff			x				x
Accuracy		x	x				
Detail tie-in	x	x					
Completeness		x	x				
Classification			x			x	
Rights			x	x			
Presentation and Disclosure				x		x	

Table 2 : Relation of A/R Audit Objectives with Substantive Audit Tests

Each test will provide assurance related to all objectives that are covered. Therefore, required assurance is computed by considering Pl(MM) with respect to those objectives. The results are given in the third column of **Table 4.**

16 Then detection risk is about 14.8 % and yields 5 % AAR when combined with CR and IR.
17 Adapted from Arens and Loebbecke (1994), p. 328.

Substantive test	Objectives covered	Pl(MM) w.r.t. objectives after IR, CR	Pl(MM) w.r.t. objectives after IR, CR, (2),(4),(6)	Pl(MM) w.r.t. objectives after IR, CR (2),(4),(6),(3)
(1)	others	0.252	0.000504	0.0000504
(2)	existence, realizable value others	0.3384	0.058608	0.0577008
(3)	existence cutoff others	0.36	0.36	0.036
(4)	others	0.252	0.000504	0.0000504
(5)	realizable value	0.288	0.0576	0.0576
(6)	existence, others	0.252	0.00504	0.000504
(7)	cutoff	0.36	0.36	0.036
Pl(MM) w.r.t. all objectives :		**0.36**	**0.36**	**0.08784**

Table 4 : Effect of Audit Tests on the Plausibility of Material Misstatement

The auditor decides to perform procedures (2), (4) and (6) on the whole population of data because of their relative low cost and because they cover objectives that are not checked in other ways. Procedure (2) results in no material misstatements and two immaterial realizable value misstatement put on the unadjusted error worksheet and a subjective bba :

$m_{(2)}(\{(x,y,z,u) \mid x,y,u \in \{C\}\}) = 0.7$

$m_{(2)}(\{(x,y,z,u) \mid x,u \in \{C\}, y \in \{IM, C\}\}) = 0.1$

Procedures (4) and (6) yield no misstatements, and the subjective bba:

$m_{(4)}(\{(x,y,z,u) \mid u \in \{C\}\}) = 0.9$

$m_{(6)}(\{(x,y,z,u) \mid x,u \in \{C\}\}) = 0.9$

The bba defined on the product set that represents the evidence from IR and CR and procedures (2),(4), (6) is included in Appendix III. Although the bba contains all information, it can be more instructive to consider the effect of these substantive tests on the plausibility of material misstatement (see column 4 in **Table 4**). The value of Pl(MM) w.r.t. the objectives covered by the procedures already performed, decreases strongly. This is to be expected as it is the purpose of performing these tests. In addition, we see that for procedures with overlapping objectives, Pl(MM) w.r.t. to the objectives they cover has also decreased. Hence, when planning these remaining procedures, this new value of Pl(MM) should influence required assurance, because part of the assurance has been provided by procedures with overlapping objectives. Finally, it is clear that Pl(MM) with respect to cutoff (to be tested by (7)) has not decreased because none of the procedures already performed provided evidence on cutoff. The overall Pl(MM) and Pl(MM) related to (3) also remain at 0.36 because cutoff has not been checked yet.

The auditor can now proceed testing. Using **Table 4** he decides to perform (3) and sets the required assurance at 0.75. This level of assurance would provide at most 10 % plausibility of material misstatement after testing, because 0.25 x 0.36 < 0.10. Using MUS, he computes the appropriate sample size and performs the tests. All differences were reconcilable except for 3 immaterial accuracy misstatements. The resulting bba is :

m({(x,y,z,u) | x,z,u ∈ {C}}) = 0.7

m({(x,y,z,u) | x,z ∈ {C}, u ∈ {IM,C}}) = 0.2

The assurance provided is 0.9 and therefore meets the requirement. Combining with previous evidence, leads to the combined bba that is included in Appendix IV.

Again, we will interpret the effect of this additional evidence by looking at the reduced plausibilities in column 5 of **Table 4**. Overall Pl(MM) is reduced to about 8.7 %. When this value is still too high, the auditor can continue evidence collection. The most effective way to proceed is by performing procedure (5) because both highest plausibilities in column 5 are associated with tests on realizable value. The fact that realizable value is the best candidate for continued testing, can also be found using the projection of the combined belief function to individual objectives frames, yielding :

Pl(MM) w.r.t. existence = 0.000504 | Pl(MM) w.r.t. cutoff = 0.036

Pl(MM) w.r.t. realizable value = 0.0576 | Pl(MM) w.r.t. others = 0.0000504

When the overall assurance is sufficient, the auditor can perform remaining tests for only minimal sample sizes, or can stop evidence collection. In this illustration, we will assume that the overall Pl(MM) of 8.7 % is acceptable. Hence, evidence collection at the lowest level of the framework is completed.

Completing the Audit of the Sales and Collection Cycle

The opinion on the cycle as a whole can be obtained by aggregation using the incomplete refinement between L3 and L2 in **Figure 1.** The transfer of mass from the product set to the cycle frame SC, leads to a bba at the cycle level equal to :

m({C}) = 0.63

m(SC) = 0.37

The mass committed to the frame SC, is made up of two parts, being :

- the part that is temporarily associated with the frame because it is not yet known whether combined immaterial errors will become material (0.28216)
- the part that is definitely committed to the frame because the subsets from which mass is transferred contain MM, IM and C values for major objectives (0.08784).

At this stage, the unadjusted error sheet is evaluated to judge whether the combined effect of immaterial misstatements exceeds tolerable error. Three cases are possible :

Case I : When combined error is immaterial, the subset of hypotheses associated with this situation and supported by a mass of 0.28216, can be labeled {C, IM} at the cycle level. The resulting bba is then : m({C}) = 0.63; m({C, IM) = 0.28216; m({SC}) = 0.08784

Case II : When combined error is material, the auditor will ask for adjustment. Via adjustment, the situation (supported by a mass of 0.28216) is corrected in such a way that the combined effect is no longer material. The mass can then also be shifted to {C, IM} which yields the same bba as in Case I.

Case III : It is possible that the client does not agree with the adjustments asked for in case of material combined error. When no reconsideration of tolerable error is made, then the auditor must conclude that this situation (supported by a mass of 0.28216) leads to material misstatement and will shift this mass to {MM} : m({C}) = 0.63 ; m({MM) = 0.28216 ; m({SC}) = 0.08784

Again, it is instructive to consider the value of Pl(MM), Bel(MM), Bel(C) and Bel(IM,C) in different cases (see **Table 5**) and compare it with their values before evaluating the unadjusted error sheet.

	Before unadjusted error sheet evaluation	**After unadjusted error sheet evaluation**		
		Case I	**Case II**	**Case III**
Pl(MM)	0.37	0.08784	0.08784	0.37
Bel(MM)	0	0	0	0.28216
Bel(C)	0.63	0.63	0.63	0.63
Bel(IM, C)	0.63	0.91216	0.91216	0.63

Table 5 : Effect of Unadjusted Error Sheet Evaluation at the Cycle Level

Before considering the combined effect of immaterial errors on the fairness of the cycle, Pl(MM) is 0.37, because at that stage material misstatement can still be concluded for these errors, whereas Bel(MM) is zero because no evidence in favor of MM has yet been found. Belief in {IM, C} equals belief in {C} because there is also no evidence for immaterial errors to remain immaterial.

After evaluation, Case I and Case II lead to the same situation because the final result is the judgment that errors (after possible adjustment) are immaterial. This lowers Pl(MM) for the cycle to about 8.7 % which equals Pl(MM) for all objectives at the lower level[18]. Case III leads to support for {MM} and therefore Bel(MM) greater than zero.

Combination of Evidence for several Cycles

We continue this simplified illustration by assuming that the audit of a second cycle (C2) is completed and no other cycles are audited. The result of the second cycle is described by the bba : m({C}) = 0.7 ; m({C, IM}) = 0.25 ; m(C, IM, MM) = 0.05. Furthermore, we assume for SC the bba described in Case I above.

In order to combine the bba's from both cycles, they will be extended to the frame, SC X C2, next combined with Dempster's rule and finally propagated upwards to FS using the incomplete refining (computational details are included in Appendix V).

The resulting bba defined on FS equals:

m({C}) = 0.441
m(FS) = 0.559

The largest part of the mass committed to the frame consists of a temporary commitment (equal to 0.425552) because it is not decided yet whether combination of immaterial cycle errors will lead to material misstatement of the FS. The remainder is the part that is definitely committed to the frame (0.133448). Again, combined error on the unadjusted error

[18] In general, Pl(MM) at L2 after judging combined error immaterial is smaller than or equal to Pl(MM) w.r.t. all objectives at L3.

sheets will be evaluated and compared with materiality. In case the combined effect is material and no adjustment is made, the mass equal to 0.425552 will be shifted from the frame to {MM}; else the mass is shifted to {C, IM}. Belief and plausibility values in both cases are given in **Table 6**.

	Before unadjusted error sheet evaluation	After unadjusted error sheet evaluation - Case I for SC combined with C2	
		Material effect-unadjusted	Material effect - adjusted, or immaterial effect
Pl(MM)	0.559	0.559	0.133448
Bel(MM)	0	0.425552	0
Bel(C)	0.441	0.441	0.441
Bel(IM, C)	0.441	0.441	0.866552

Table 6 : Effect of Unadjusted Error Sheet Evaluation at the FS level

It is obvious that the auditor will issue an adverse opinion in case of an unadjusted material effect on the FS as a whole. When the effect is material but adjusted or it is immaterial, the auditor may accept the Pl(MM) and judge whether a qualification of his opinion is necessary because of the substantial difference between Bel(C) and Bel(IM, C). As part of completing the audit, analytical procedures will be performed that provide evidence at the level of FS and can lower the value of Pl(MM). If insufficient, the auditor can also return to lower levels and expand detailed testing in order to decrease Pl(MM).

Conclusions

At the top level, the most important measure is Pl(MM), because it quantifies the risk that the auditor is taking when the conclusion "free of material misstatements beyond reasonable doubt" is issued. In addition, the auditor will be interested to know whether there is also evidence in favor of {MM}. This information is contained in Bel({MM}). When there is substantial evidence, the issuance of an adverse opinion is the most appropriate action. Furthermore there is interesting information in Bel({C}) and Bel({IM, C}) because it quantifies the support for concluding the fairness of FSs. At lower levels, Pl(MM) for groups of variables (objectives, cycles) is of particular interest. When evidence bears on several variables (objectives) at the same time, it is useful to know what the combined assurance of testing should be in order to meet required assurance.

The most common scenario was illustrated. Because material misstatements are almost always corrected, mass will seldom be committed to {MM}. If it is, then conflict will arise that will not be normalized but will be propagated to conflict at the top level in case of outer conflict[19], or will disappear in case of inner conflict. When conflict appears at the top level, then there was support for {MM}, because due to the way of representing evidence, it cannot arise in another way. Hence, conflict means support for material misstatements.

Discussion

In comparison with the ARM, the framework's greatest advantages are the representation of the intended meaning of audit risks, the integration of evidence at different levels and the availability of risk measures for groups of objectives. With respect to representing audit risks, the intended meaning of inherent and control risk as the possibility of material misstatement can be modeled by Pl(MM) which is different from Bel(MM), or the necessity of material misstatement. In the ARM, possibility is represented in the same way as necessity. With respect to integrating evidence, the framework provides rules of reasoning that allow conclusions at other levels than those at which evidence was initially represented. The ARM is a "one-item-model" : it can be used for the FS as a whole, it can be used for a particular account in separation, it can be used for a particular cycle in separation etc., but there are no links to go from e.g. the accounts of a cycle to the cycle itself, even though it is common knowledge that fairness of all accounts in a particular cycle must allow conclusions to be drawn for the cycle as a whole. Caused by the ARM's focus on one item at the time, is the lack of measures of assurance or risk related to tests that cover more than one item. As a consequence, the ARM is not really suited to plan such tests. The framework can provide measures of risk or assurance for whatever combination of items that is needed for planning efficient testing. When needed, it can be extended with more accounts per cycle or transactions per account according to the same principles.

In comparison with Srivastava's model, the framework presented here deals with three major limitations. Firstly, the use of AND-relationships at all levels in Srivastava's model implies that the only reasoning that is possible is the conclusion of fairness at a higher level if and only if all composing elements at the lower level are considered fair. In our structure, the use of refinings allows for the flexibility of defining whatever relation between detailed and aggregate hypotheses. Related to this issue is Srivastava's questioning of the appropriateness of the same relation in both directions of the framework. The use of the different refinings in upward and downward transfer in our framework, solves this problem. The second limitation discussed by Srivastava is the definition of frames with only two hypotheses (materially misstated and correct) which does not allow for immaterial errors to become material after combination. The inclusion of the hypothesis "immaterially misstated" and the definition of the incomplete refining provide the possibility to keep track of the unadjusted immaterial misstatements and their combined effect. Finally, in Srivastava's structure evidence affecting more than one variable at the time cannot directly be represented. Because audit tests almost always cover more than one variable (objective), our structure includes products sets that allow for the representation of evidence bearing on more than one objective.

To use our framework for real life applications, software is needed to perform evidence combination and propagation and to compute measures of interest to the auditor. As the

[19] Outer conflict is present when the conflict would have arisen in the coarsening as well. Inner conflict is present when it would not have occurred when evidence were to be presented in the coarsening.

amount of information available is enormous, it would be sensible to use such a system by questioning it instead of having it produce all possible measures of belief and plausibility. Depending on the stage at which questioning is performed, the framework can be used for evaluation is well as for planning purposes.

Appendix I : Deriving belief from statistical evidence

The method used in this framework is built on two assumption, being (1) proportionality between the probability of x given θ, $p_\theta(x)$, and $Pl_x(\theta)$ and (2) consonance, and is described formally as :

$Bel_x : 2^\Theta \rightarrow [0, 1]$ is a consonant belief function with n focal elements F of the form :

$$F_0 = \bigcup_{p_\theta(x)=\max_{\varepsilon\in\Theta}\{p_\varepsilon(x)\}} \{\theta\}$$

$$F_j = F_{j-1} \cup \{ \bigcup_{p_\theta(x)=\max_{\varepsilon\in\Theta\setminus F_{j-1}}\{p_\varepsilon(x)\}} \{\theta\} \}$$

with bpa :

$$m_x(F_0) = \frac{\sum_{\theta\in F_0} p_\theta(x)}{\sum_{\theta\in\Theta} p_\theta(x)}$$

$$m_x(F_j) = \frac{\sum_{\theta\in F_j\setminus F_{j-1}} p_\theta(x)}{\sum_{\theta\in\Theta} p_\theta(x)}$$

and belief for focal elements :

$$Bel_x(F_0) = \frac{\sum_{\theta\in F_0} p_\theta(x)}{\sum_{\theta\in\Theta} p_\theta(x)}$$

$$Bel_x(F_j) = \frac{\sum_{\theta\in F_j} p_\theta(x)}{\sum_{\theta\in\Theta} p_\theta(x)}$$

and for subsets other than focal elements :

$$Bel_x(A) = \underset{F_{j-1}\subset A\not\subset F_j}{Bel_x}(F_{j-1})$$

The simplicity of this method is caused by the characteristic that the belief committed to a subset A of Θ equals the belief committed to its largest subset that is also a focal element. This means that all elements θ of A not contained in that largest subset do not have any contribution to the total belief in A, meaning that the information contained in $p_\theta(x)$ for these elements is ignored when computing the belief in A.

This reasoning of ignoring information (or selectively withholding information) is inspired by the simplifying way in which statistical evidence is often used in auditing practice: the sample outcome x is used to assess the probability $p_\theta(x)$ for each θ that is a relevant population value or range of population values. Next, only the value of θ for which $p_\theta(x)$ is maximal, is retained as the population value supported by the sample evidence. This corresponds to the situation in which a decision about the single most likely hypothesis would be required. Knowledge of the probability of the outcome x given other hypotheses, is discarded. In our approach, the number of focal elements is extended gradually, with F_n representing the result of the decision in auditing practice in a situation in which the n most likely hypotheses would be required. In that way, statistical information that is incorporated in the belief function exceeds information retained in auditing practice.

Appendix II : Substantive tests for A/R

(1) Obtain an aged list of receivables : trace accounts to the master file, foot schedule and trace to general ledger.
(2) Obtain an analysis of the allowance for doubtful accounts and bad debt expense : test accuracy, examine authorization for write-offs, and trace to general ledger.
(3) Obtain direct confirmation of accounts receivable and perform alternative procedures for nonresponses.
(4) Review receivables for any that have been assigned or discounted.
(5) Investigate collectibility of account balances.
(6) Review lists of balances for amounts due from related parties or employees, credit balances, and unusual items, as well as notes receivable due after one year.
(7) Determine that proper cutoff procedures were applied at the balance sheet date to ensure that sales, cash receipts, and credit memos have been recorded in the correct period.

Appendix III : Combined bba for IR and CR and procedures (2),(4) and (6)

$m(\{(x,y,z,u) \mid x,y,u \in \{C\}, z \in \{IM,C\}\}) = 0.448$
$m(\{(x,y,z,u) \mid x,y,u \in \{C\}\}) = 0.252$
$m(\{(x,y,z,u) \mid x,u \in \{C\}, y,z \in \{IM,C\}\}) = 0.1792$
$m(\{(x,y,z,u) \mid x,u \in \{C\}, y \in \{IM,C\}\}) = 0.04896$
$m(\{(x,y,z,u) \mid x,u \in \{C\}\}) = 0.05184$
$m(\{(x,y,z,u) \mid u \in \{C\}, x,y,z \in \{IM,C\}\}) = 0.01152$
$m(\{(x,y,z,u) \mid u \in \{C\}, x,y \in \{IM,C\}\}) = 0.000389$
$m(\{(x,y,z,u) \mid u \in \{C\}, x \in \{IM,C\}\}) = 0.001555$
$m(\{(x,y,z,u) \mid u \in \{C\}, y \in \{IM,C\}\}) = 0.000907$
$m(\{(x,y,z,u) \mid u \in \{C\}\}) = 0.003629$
$m(\{(x,y,z,u) \mid x,y,z,u \in \{IM,C\}\}) = 0.00128$
$m(\{(x,y,z,u) \mid x,y,u \in \{IM,C\}\}) = 0.000043$
$m(\{(x,y,z,u) \mid x,u \in \{IM,C\}\}) = 0.000173$
$m(\{(x,y,z,u) \mid y \in \{IM,C\}\}) = 0.000101$
$m(\Theta) = 0.000403$

Appendix IV : Combined bba for IR and CR and procedures (2),(4),(6) and (3)

$m(\{(x,y,z,u) \mid x,y,u \in \{C\}, z \in \{IM,C\}\}) = 0.0448$
$m(\{(x,y,z,u) \mid x,y,u \in \{C\}\}) = 0.0252$
$m(\{(x,y,z,u) \mid x,u \in \{C\}, y,z \in \{IM,C\}\}) = 0.01792$
$m(\{(x,y,z,u) \mid x,u \in \{C\}, y \in \{IM,C\}\}) = 0.004896$
$m(\{(x,y,z,u) \mid x,u \in \{C\}\}) = 0.005184$
$m(\{(x,y,z,u) \mid u \in \{C\}, x,y,z \in \{IM,C\}\}) = 0.001152$
$m(\{(x,y,z,u) \mid u \in \{C\}, x,y \in \{IM,C\}\}) = 0.0000389$
$m(\{(x,y,z,u) \mid u \in \{C\}, x \in \{IM,C\}\}) = 0.0001555$
$m(\{(x,y,z,u) \mid u \in \{C\}, y \in \{IM,C\}\}) = 0.0000907$
$m(\{(x,y,z,u) \mid u \in \{C\}\}) = 0.0003629$
$m(\{(x,y,z,u) \mid x,y,z,u \in \{IM,C\}\}) = 0.000128$
$m(\{(x,y,z,u) \mid x,y,u \in \{IM,C\}\}) = 0.0000043$
$m(\{(x,y,z,u) \mid x,u \in \{IM,C\}\}) = 0.0000173$
$m(\{(x,y,z,u) \mid y \in \{IM,C\}\}) = 0.0000101$
$m(\{(x,y,z,u) \mid x,y,z,u \in \{C\}\}) = 0.63$
$m(\{(x,y,z,u) \mid x,z,u \in \{C\}, y \in \{IM, C\}\}) = 0.2178752$
$m(\{(x,y,z,u) \mid x,z,u \in \{C\}\}) = 0.0517248$
$m(\{(x,y,z,u) \mid x,z \in \{C\}, y,u \in \{IM,C\}\}) = 0.0002848$
$m(\{(x,y,z,u) \mid x,z \in \{C\}, u \in \{IM,C\}\}) =0.0001152$
$m(\Theta) = 0.0000403$

Appendix V : Cylindrical extensions and combination of the bba's of SC and C2

The cylindrical extension to the frame, SC X C2, yields :

$m_{SC}(\{(x,y) \mid x \in \{C\}\}) = 0.63$
$m_{SC}(\{(x,y) \mid x \in \{IM, C\}\}) = 0.28216$
$m_{SC}(SC \times C2) = 0.08784$
$m_{C2}(\{(x,y) \mid y \in \{C\}\}) = 0.7$
$m_{C2}(\{(x,y) \mid y \in \{IM, C\}\}) = 0.25$
$m_{C2}(SC \times C2) = 0.05$

Combination with Dempster's rule then yields :

$m(\{(x,y) \mid x,y \in \{C\}\}) = 0.441$
$m(\{(x,y) \mid x, y \in \{IM, C\}\}) = 0.07054$

$m(\{(x,y) \mid x \in \{C\}, y \in \{IM, C\}\}) = 0.1575$

$m(\{(x,y) \mid x \in \{C\}\}) = 0.0315$

$m(\{(x,y) \mid x \in \{IM, C\}, y \in \{C\}\}) = 0.197512$

$m(\{(x,y) \mid x \in \{IM, C\}\}) = 0.014108$

$m(\{(x,y) \mid y \in \{C\}\}) = 0.061488$

$m(\{(x,y) \mid y \in \{IM, C\}\}) = 0.02196$

$m(SC \times C2) = 0.004392$

Next, mass is propagated upwards from the product set of cycles to the FS-frame using the incomplete refining, leading to a bba defined on FS equal to :

$m(\{C\}) = 0.441$
$m(FS) = 0.559$