

Author's Accepted Manuscript

Activity labeling in process modeling: empirical insights and recommendations

J. Mendling, H.A. Reijers, J. Recker

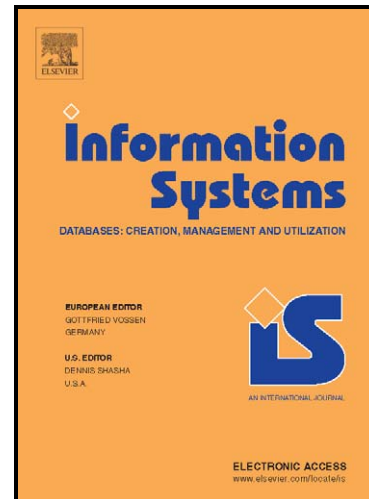
PII: S0306-4379(09)00031-3
DOI: doi:10.1016/j.is.2009.03.009
Reference: IS 611

To appear in: *Information Systems*

Received date: 30 April 2008
Revised date: 16 March 2009
Accepted date: 22 March 2009

Cite this article as: J. Mendling, H.A. Reijers and J. Recker, Activity labeling in process modeling: empirical insights and recommendations, *Information Systems* (2009), doi:[10.1016/j.is.2009.03.009](https://doi.org/10.1016/j.is.2009.03.009)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



www.elsevier.com/locate/infosys

Activity Labeling in Process Modeling: Empirical Insights and Recommendations

J. Mendling^{a,*} H.A. Reijers^b J. Recker^c

^a*Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany*

^b*Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven,
The Netherlands*

^c*Queensland University of Technology, 126 Margaret Street, QLD 4000 Brisbane,
Australia*

Abstract

Few studies have investigated the factors contributing to the successful practice of process modeling. In particular, studies that contribute to the act of developing process models that facilitate communication and understanding are scarce. Although the value of process models is not only dependent on the choice of graphical constructs but also on their annotation with textual labels, there has been hardly any work on the quality of these labels. Accordingly, the research presented in this paper examines activity labeling practices in process modeling. Based on empirical data from process modeling practice, we identify and discuss different labeling styles and their use in process modeling practice. We perform a grammatical analysis of these styles and use data from an experiment with process modelers to examine a range of hypotheses about the usability of the different styles. Based on our findings, we suggest specific programs of research towards better tool support for labeling practices. Our research contributes to the emerging stream of research investigating the practice of process modeling and thereby contributes to the overall body of knowledge

about conceptual modeling quality.

Key words: Business Process Modeling, Model Quality, Survey, Systems Analysis and Design;

1 Introduction

In recent years, the conceptual mapping of processes in the form of process models has emerged as a primary reason to engage in conceptual modeling [1] and is considered as a key instrument for the analysis and design of process-aware information systems [2], service-oriented architectures [3], and web services [4] alike. To that end, process models typically describe in a graphical way at least the activities, events, states, and control flow logic that constitute a business process [5]. Additionally, process models may also include information regarding the involved data, organizational and IT resources, and potentially other artifacts such as external stakeholders and performance metrics, see e.g. [6]. Similar to other forms of conceptual modeling, process models are first and foremost required to be intuitive and easily understandable, especially in information systems project phases that are concerned with requirements documentation and communication [7].

Process modeling has been around for some thirty years. However, only of late has research started to examine quality aspects pertaining to process modeling. In fact, quality issues of conceptual modeling in general have only

* Corresponding author

Email addresses: jan.mendling@wiwi.hu-berlin.de (J. Mendling),
h.a.reijers@tue.nl (H.A. Reijers), j.recker@qut.edu.au (J. Recker).

18 recently been receiving increased attention in academia [8]. Notwithstanding
19 the research findings collected to date, surprisingly little is known about the
20 actual “practice of process modeling” and the factors that contribute to build-
21 ing a “good” process model, for example one that aids human understanding
22 of the depicted business domain [9]. Work has been carried out, for instance,
23 that examined the impact of process model structure, model user competency
24 and process modeling language on process model understanding. While the
25 impact of structural properties is clearly identified [10], it is also reported
26 that model readers systematically overestimate their ability to draw correct
27 conclusions from a model [9]. It was also found that the choice of languages
28 used for process modeling (e.g., BPMN versus EPCs) has only insignificant
29 effects on process model understanding [11]. Other research has successfully
30 investigated the graphical constructs and their meaning in process models,
31 e.g., [12], the expressiveness and validity of control flow aspects in process
32 models, e.g., [13], or process-related aspects such as data and resources, e.g.,
33 [14,15].

34 This situation raises the question of other antecedents of process model un-
35 derstandability. Most of the previous work has focused on syntactic quality
36 aspects [16]. In contrast, semantic and pragmatic aspects of model quality have
37 mostly been neglected. In particular, little attention has been devoted to a very
38 essential task in process modeling - the *labeling* of the graphical constructs,
39 in particular of the constructs representing “activities” (or “tasks”, or “work
40 to be performed”) in a process model. This is rather surprising given that –
41 clearly – the true meaning of any construct in a process model is only revealed
42 when model users read and *intuitively* understand the labels assigned to the
43 construct. Current practice indicates that the labeling of activity constructs

44 is a rather arbitrary task in modeling initiatives and one that is sometimes
45 done without a great deal of thought [17]. This can undermine the under-
46 standability of the resulting models in cases where the meaning of the labels
47 is ambiguous, not readily understandable, or simply counter-intuitive to the
48 reader.

49 Accordingly, in our work we seek to address this gap and contribute to the
50 existing line of work towards more understandable process models. The ob-
51 jective of our research is to investigate the styles that are in use to annotate
52 activities in process models and how these styles affect the understandability
53 of such models.¹ More precisely, the *aim of this paper* is to suggest, based
54 on our empirical findings, an imperative style for modelers to create more
55 understandable process models.

56 We proceed as follows. In Section 2 we discuss the theoretical foundation for
57 our work and investigate current labeling practices in process modeling. In
58 Section 3 we discuss design of, conduct of, and findings from an experiment
59 with process modelers. In Section 4 we then discuss the implications of our
60 findings and suggest specific programs of research towards better support for
61 process model labeling practices. We conclude in Section 5 by reviewing our
62 contributions, and discussing some conclusions.

¹ We recognize the need to extrapolate our research to other aspects of process models, such as the data, resource and control flow perspective. We deemed the focus on ‘activity constructs’ a suitable starting point for our endeavor due to the centrality of the ‘activity’ concept in process modeling.

63 2 Background

64 In presenting the background to our research, we refer to a theory of multi-
65 media learning originating from cognitive science. This theory suggests that
66 labeling practices are indeed significant factors contributing to how well or
67 how poorly process models can be understood by their end users. To deter-
68 mine what a good labeling style is, we then identify different styles of labeling
69 being used in practice. We describe how the exploration of a large number
70 of real-life process models gives us this insight. One of the styles that is en-
71 countered is the usage of verb-object labels. As this style is widely promoted
72 in the literature [18,19,20], we formulate several hypotheses on its presumed
73 superiority over the other styles encountered in our exploration.

74 2.1 Theoretical Foundation

75 Dual Coding Theory [21] suggests that individuals have two separate chan-
76 nels – visual and auditory – that they use when processing information. The
77 two channels complement each other, such that receiving simultaneous infor-
78 mation through each channel improves understanding compared to receiving
79 information through one channel only. In other words, individuals understand
80 informational material better when it is provided through both auditory (i.e.,
81 words) and visual (i.e., images) channels.²

82 Based on this observation, the Cognitive Theory of Multimedia Learning
83 (CTML) [23,24] suggests that learning material intended to be received, un-

² Indeed, most people read by speaking out the words of the text in their mind, which even suppresses visual activation [22].

84 derstood and retained by its recipients should be presented using *both* words
85 and pictures. This sounds conducive to the task of process modeling, where
86 both visual (graphical constructs) and auditory (labels and text annotations)
87 material are available to add information about a business domain in a pro-
88 cess model. However, due to the overall limited number of graphical constructs
89 used in a process model – there are typically few if not only one graphical con-
90 struct for representing activities – most of the critical domain information is
91 contained in the textual labels of the constructs, viz., in *auditory* channels.
92 Based on CTML it can thus be expected that model understanding can be
93 improved if better guidance can be provided for the act of labeling of process
94 model constructs.

95 The general principle that our expectation builds on is described by Mayer
96 [24] as the “Multimedia Principal”. And indeed, prior research on conceptual
97 modeling has successfully demonstrated that the multimedia principal informs
98 model understanding. Empirically observable differences in model understand-
99 ing based on the multimedia principal were found, for instance, in the data
100 modeling domain [25,26] as well as in the process modeling domain [11].

101 2.2 Labeling Styles in Practice

102 For business process modeling, the labeling of constructs such as activities
103 is often more art than science. In practice, a number of informal guidelines
104 exist that typically suggest a verb-object convention (e.g., “approve order”,
105 “verify invoice”) for labeling activities, e.g., [18,19,20]. This convention is sim-
106 ilar to a style that is advocated in guidelines that support the creation of
107 understandable use case descriptions, a widely accepted requirements tool in

108 object-oriented software engineering [27,28]. We will refer to this labeling style
109 of activities as the *Verb-Object Style*. But as much promotion it receives in the
110 process modeling domain, both anecdotal evidence and causal inspection of
111 real process models indicate that this labeling style is neither universally nor
112 consistently applied. Even the practical guide for process modeling with ARIS
113 [29, pp.66-70] shows models with both actions as verbs and as nouns. Also,
114 one may think that the more information contained in the labels, the clearer
115 the meaning will be to the reader. Recent research, however, uncovered that
116 shorter activity labels improve model understanding [30].

117 To get a better idea of the variety in labeling styles being applied in practice,
118 we turn to the SAP Reference Model [31]. The development of the SAP refer-
119 ence model started in 1992 and first models were presented at CEBIT'93 [31,
120 p.VII]. Since then, it was developed further until version 4.6 of SAP R/3, which
121 was released in 2000. Overall, the SAP reference model includes 604 business
122 process models depicted using the Event-driven Process Chains (EPC) nota-
123 tion, capturing information about the SAP R/3 functionality to support the
124 business processes in a wide range of organizations. With the SAP solution be-
125 ing the market leading tool in the Enterprise Systems market we feel that the
126 examination of SAP process models gives us a good understanding of the use of
127 process models in real-life business contexts. Amongst other application areas,
128 the SAP reference model denotes a frequently used tool in the implementation
129 of SAP systems [32], and much literature has covered its development and use
130 [31]. Furthermore, it is frequently referenced in research papers as a typical
131 reference model and used in previous examinations of process modeling, e.g.,
132 [10,33,34].

133 Altogether, the 604 EPC models in the SAP reference model include 19,838

134 activity labels, which we all manually inspected and classified. In 94% of these
135 cases (18,648 instances), the activity labels refer to a certain action that should
136 be undertaken, such as *Check billing block* or *Order Execution*. This is not so
137 for 6% of the labels, because they neither include a verb nor a noun that refers
138 to an action, consider, for instance, “Status Analysis Cash Position”. We will
139 refer to this style as the *Rest* category.

140 Note that the EPC models considered were designed based on the functionality
141 and the terminology of the SAP system which might create different biases. On
142 the one hand, system terminology could potentially be less intuitive compared
143 to labeling in conceptual design models. On the other hand, the labels could
144 be more precise than labels in conceptual modeling practice. Yet, neither the
145 high frequency of verb-object styles nor the variety of labeling styles in use
146 directly suggest such bias.

147 Despite the wide proliferation of 18,648 “action-oriented” labels of the 19,838
148 activity labels in the SAP reference model overall, this situation does not imply
149 that the verb-object style is strictly enforced within this subset. Rather, it is
150 applied to only about two third of the “action-oriented” labels (60 % of *all*
151 activity labels). The remaining subset of the “action-oriented” labels (34 % of
152 *all* activity labels) denote labels where the action is grammatically captured
153 *as a noun*. This noun can be either a gerund of the verb or a noun that is
154 derived from a verb, like *Order Processing* or *Invoice Verification*. We will
155 refer to this style of labeling as the *Action-Noun Style*. The overall result from
156 classifying all 19,838 activity labels can be seen in Table 1.

157 We will now consider this data in more detail. More precisely, for each of
158 the labeling styles found, we perform a grammatical analysis using the lexical

| <i>Verb-Object Labels</i> | <i>Action-Noun Labels</i> | <i>Rest</i> | <i>Sum</i> |
|---------------------------|---------------------------|-------------|------------|
| 11,830 | 6,808 | 1,201 | 19,838 |
| 60% | 34% | 6% | 100% |

Table 1

Distribution of Activity Label Styles in the SAP Reference Model

159 database WordNet [35] to identify potential types of interpretation ambiguity.
 160 This grammatical analysis builds on the identification of syntactic categories
 161 such as *noun* and *verb*. Further categories like *adjective* and *adverb* could also
 162 be used but do not pertain to activity labeling in process modeling, which
 163 is why we excluded these categories from our analysis. For many words, the
 164 syntactic category can be identified purely syntactically, as for instance with
 165 the word *grammar*, which is a noun. Some words, however, are *ambiguous* re-
 166 garding the category they belong to (when analyzed in isolation). Consider the
 167 word *design*, which can be a verb (*to design*) or a noun (*the design*) depending
 168 on the grammatical context. As these examples from natural language process-
 169 ing show, ambiguity can be a significant impediment to ease of understanding.
 170 In light of this observation we thus argue that those labeling styles should be
 171 considered in process modeling that are least susceptible to ambiguity. We
 172 illustrate our argument with examples from the SAP Reference Model:

173 **Verb-object labels.** Most of the verb-object labels seem intuitively under-
 174 standable to us. Still, there are some cases that are ambiguous from a gram-
 175 matical point of view: The English language allows for a so-called *zero deriva-*
 176 *tion* beyond the suffix *-ize* and the suffix *(i)fy* derivation of verbs from nouns
 177 [36]. As a consequence, the same word can both be a noun and a verb. Con-
 178 sider, for example, the labels *Measure Processing*, *Export License Check*, and

179 *Process Cost Planning*. They have in common that the first word can be a
180 verb, but reading it as an object describing an action is also possible. *Measure*
181 *Processing* could potentially refer to the processing of a measure or to the
182 measurement of a processing. The same observation holds for the other labels.
183 Some of these ambiguities can be resolved by considering context information
184 such as the labels of the other activities in the same process model. If the
185 verb-object style was consistently used as a standard throughout a process
186 model, it would be clear to interpret the first term as a verb.

187 **Action-noun labels.** With respect to action-noun labels, some of these can
188 be easily interpreted, but again there can be cases of grammatical ambiguity.
189 Consider, for instance, *Notification Printing*. Again, there are two potential
190 interpretations: a notification is printed, or someone is notified of a printing
191 job. Alternatively, the verb could just have been forgotten by the modeler.
192 This interpretation is likely in cases where the action noun could also be an
193 object, like *Order*, which can refer to both an action or an object. We call this
194 type of ambiguity the *action-object ambiguity*. In such cases, the model reader
195 might be tempted to infer the action by considering the context of the activity.
196 Syntactically, the label could be easily extended with such semantically diverse
197 verbs as *start*, *stop*, or *schedule*. Using a verb-object style would have avoided
198 the problem of action-object ambiguity and the necessity of having to infer a
199 verb to establish the appropriate meaning.

200 **Rest labels.** Some of the rest labels clearly point to a specific business object,
201 for instance *Status Analysis Cash Position*, such that a verb could potentially
202 be inferred from the context. Yet there are also activity labels like *DEÜV*
203 and *Jamsostek* that are altogether difficult to understand. Presumably, the
204 first one refers to the German regulation for data storage and transmission

205 (DEÜV Datenerfassungs- und Übertragungsverordnung) and the second to
206 the Indonesian social security system. Clearly labels of the “rest” category
207 require crystal clear context information, otherwise an inference of the action
208 to be performed is a highly problematic task due to the occurrence of *verb-*
209 *inference ambiguity*, i.e., the problem of inferring from the context of the label
210 the type of action to be performed as part of the considered process task.

211 In conclusion, the three different classes exhibit different types of ambiguities.
212 For the verb-object style, we found instances of *zero-derivation ambiguity* in
213 the SAP reference model. Altogether, we identified exactly 600 labels with
214 such ambiguity; these labels contained 23 different verbs including *change*,
215 *design*, *process*, and *report*. For the action-noun style, this problem class is
216 relevant, too. Furthermore, this style is susceptible to *action-object ambiguity*,
217 if an action noun can also refer to an object. We counted 615 cases of such
218 ambiguities. Finally, the rest group of labels, which do not mention an action at
219 all, faces *verb-inference ambiguity* (1190 cases). These three ambiguity classes
220 differ in occurrence frequency: While the *zero-derivation ambiguity* requires
221 the unlikely combination of a verb and an action object, the *action-object*
222 *ambiguity* is found more often since many documents in a business context
223 are synonymous to an action noun (e.g., *order*, *receipt*, *confirmation*). The
224 *verb-inference ambiguity* is the most significant one, since all labels of the rest
225 group suffer from it.

226 2.3 Hypotheses

227 On basis of the findings discussed above, our contention is to conjecture about
228 the influence of choice of labeling styles on the pragmatic quality of process

229 models in terms of unambiguously facilitating action [16] and usage [37]. We
230 summarize our expectations as follows. First, we formulated and grounded our
231 expectation that model understanding can be improved by guiding the act of
232 labeling following the theory of multimedia learning. In search for candidate
233 guidelines for labeling activities, anecdotal evidence, the study of the SAP
234 reference model, and our literature review suggest the verb-object labeling
235 style to be the strongest candidate style. Our empirical exploration of the
236 SAP reference model indeed confirmed the wide application of this style in
237 practice. Yet, we also found that this style is not the only style being applied:
238 A large fraction of activity labels follows an action-noun style, and there are
239 also other (“rest”) styles to be found in process models. Our grammatical
240 analysis of the three modeling styles, as described in the previous section,
241 suggested that the verb-object style appears to be the least susceptible to
242 various types of interpretation ambiguity, indicating its superiority in terms
243 of clarity of specification.

244 In light of these observations, we suggest the following primary conjecture
245 that we seek to test in our study. Based on our grammatical analysis, we
246 theorize that process modelers perceive the verb-object style to be superior to
247 the action-noun and rest labeling style alongside two dimensions:

- 248 • *Perceived ambiguity*: the degree to which an individual believes that a label
249 is ambiguous, and
- 250 • *Perceived usefulness*: the degree to which an individual believes that a label
251 is useful for understanding the process modeled.

252 This conjecture rests on the observation that the verb-objective style is less
253 prone to result in misinterpretation and confounding complexity. After all,

254 our grammatical analysis showed that it is least susceptible to ambiguity. We
255 thus advance the following two primary hypotheses we seek to test in this
256 study. First, we theorize that users working with process models have a clear
257 preference for labeling styles that avoid ambiguity:

258 **H1:** Verb-object style labels are less frequently perceived as being ambiguous,
259 followed by action-noun style labels, and finally rest labels.

260 Second, we theorize that end users working with process models have different
261 perceptions of the usefulness of the labels for understanding the process mod-
262 eled, dependent on the labeling style in which the label is articulated. More
263 specifically:

264 **H2a:** Verb-object style labels are perceived as more useful for understanding
265 the process model than action-noun style labels.

266 **H2b:** Verb-object style labels are perceived as more useful for understanding
267 the process model than rest style labels.

268 **H2c:** Action-noun style labels are perceived as more useful for understanding
269 the process model than rest style labels.

270 Hypotheses **H2a**, **H2b** and **H2c** rest on the assumption that the perceived
271 usefulness of a label is negatively influenced by the perceived ambiguity of
272 the labeling style used, based on the contention that the grammatical style
273 of a labeling type can lead to misinterpretation and confounding complexity.
274 To gather empirical evidence for this contention, we advance the following
275 additional hypothesis that we will test:

276 **H3:** Perceived ambiguity of a labeling style is negatively associated with the
277 perceived usefulness of the label.

278 In our study, we also need to consider that differences in the perceptions about
279 the ambiguity and usefulness of a process model label can also stem from differ-
280 ences between the study participants. Recent experimental research on concep-
281 tual modeling, most notably [26,38,39], has indicated significant differences in
282 the understanding of conceptual models stemming from two characteristics of
283 the conceptual model readers, these being *knowledge of the application domain*
284 (e.g., [38]) and *familiarity with the technique or notation used for conceptual*
285 *modeling* (e.g., [26]). CTML [24] suggests that previous knowledge of the do-
286 main covered in the conceptual modeling lowers the cognitive load required to
287 develop a mental model of the information displayed in the conceptual model,
288 and hence, model understanding will be easier. This is because readers can
289 bring to bear an understanding of the semantics, relevant entities or proce-
290 dures that make up the application domain depicted in a model. Similarly,
291 expertise or knowledge of the conceptual modeling artifact (i.e., the method,
292 technique or notation used) has been shown to increase the quality of the
293 models produced (e.g., [40,41]), as well as to sometimes increase the under-
294 standing of the models produced [38]. The noted interaction effects of notation
295 familiarity are speculated to stem from a modeler's self-perception about his
296 or her modeling skills. In other words, a modeler that deems himself or herself
297 to be experienced, may approach modeling tasks and outcomes differently to
298 someone that believes oneself to be a novice.

299 In light of these findings we thus advance the following, additional exploratory
300 hypotheses that seek to investigate how knowledge about the application do-
301 main and familiarity with the process modeling notation used act as moder-
302 ating variables to the propositions outlined above:

303 **H4a:** Knowledge about the application domain moderates the strength of the

304 relationship between labeling style and perceived usefulness of the label.

305 **H4b:** Familiarity with the process modeling notation moderates the strength
306 of the relationship between labeling style and perceived usefulness of the
307 label.

308 **3 Research Method**

309 *3.1 Research Design and Conduct*

310 To test the hypotheses advanced in the previous section, we developed a (self-
311 administered) questionnaire to gather quantitative insights. With this ques-
312 tionnaire we asked participants about the perceived ambiguity of certain activ-
313 ity labels, as well as their perceived usefulness. Along with the questionnaire,
314 we presented to the participant a number of activity labels as part of a spe-
315 cific process model. This has been done for several reasons. First, a label in a
316 business process model is never interpreted in isolation. Various other labels in
317 the model and the control flow relationship between the activities establish a
318 context against which a single label is interpreted. Since we do not aim to gain
319 insight into labels *per se* but in their use in process models, we have to present
320 all the labels that are discussed in the questionnaire in the context of a model.
321 Second, we had to choose a model from practice; otherwise there would have
322 been the risk that we would (unconsciously) tailor it to meet our hypotheses.
323 Third, this process model had to show a substantial variation in the labeling
324 styles being used so that we can limit potential bias in our research design.

325 Following these considerations we selected a model of a complaint process from
326 a department of a Dutch governmental agency, which is concerned with com-

327 complaint handling (see Figure 1). The model follows the EPC notation, which is
328 one of the most popular modeling techniques in industry [1]. Indeed, it is the
329 same technique as applied in the SAP reference model. In an EPC, so-called
330 *functions* (rectangles) correspond to the various tasks that may need to be ex-
331 ecuted (e.g., “Register receipt date of complaint letter”). *Events* (hexagons)
332 describe the situation before and after a function is executed (e.g., “Customer
333 at desk”). *Logical connectors* (circles) define routing rules. In particular, there
334 are three types of connectors: the logical AND for concurrency, XOR for exclu-
335 sive choices, and OR for inclusive choices. Functions, events, and connectors
336 are the classical elements of control flow modeling. These routing elements
337 are also included in other modeling languages like BPMN, YAWL, and UML
338 Activity Diagrams, which supports generalizability and repeatability of our
339 procedure.

340 The given model roughly describes the following procedure to handle the com-
341 plaints that the agency receives. A new case is opened if a new complaint is
342 received – be it by means of a phone call, personal contact, or letter. In some
343 situations, the complaint must be referred to another party, either internal
344 or external to the agency involved. Internal referrals have to be put on a so-
345 called *incident agenda*, while external referrals always require a confirmation.
346 In both cases the referral is archived in parallel. As a final step in this pro-
347 cedure, the complainant is informed. If no referral is required, a complaint
348 analysis is conducted. Later, the complaint is archived and the complainant
349 is contacted, with an optional follow up (see Figure 1).

350 The complaint process model in Figure 1 is at the heart of our questionnaire,
351 which is subdivided into three parts. In the first part we recorded demographic
352 information about the participants including gender, years of tertiary educa-

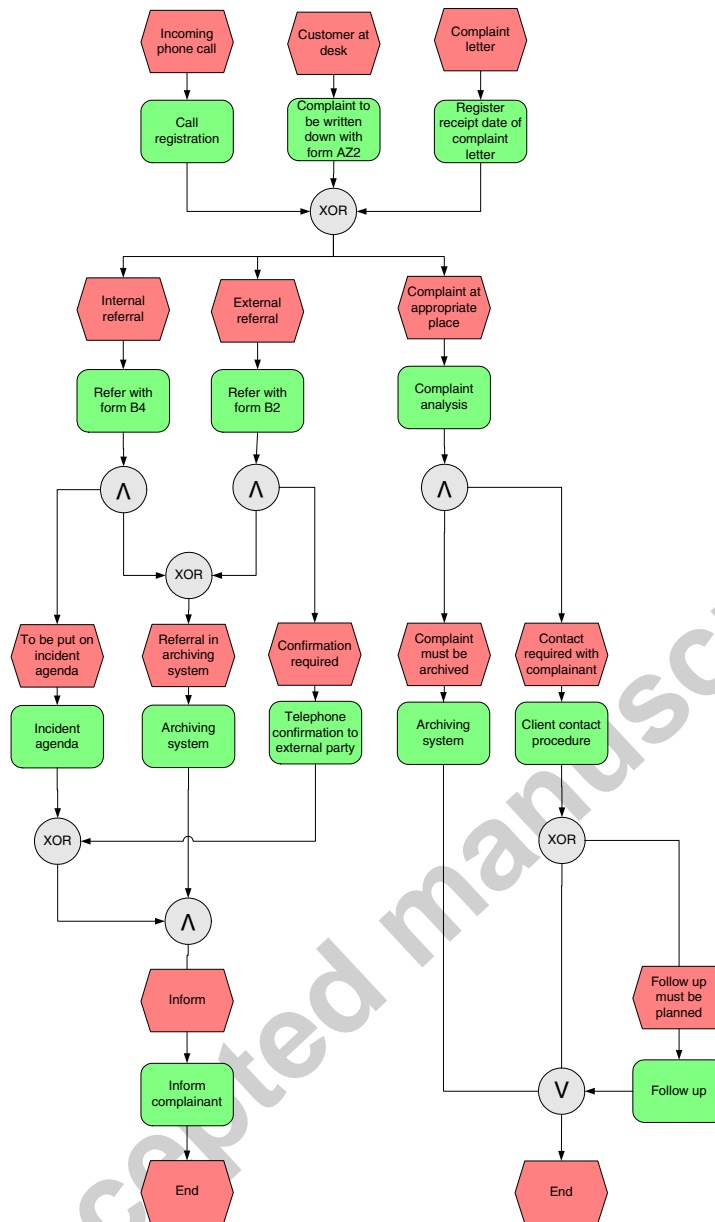


Figure 1. The complaint handling process

353 tion, preliminary knowledge of process modeling, and number of EPC models
 354 created. These questions were used to gather information about the demo-
 355 graphic distribution of the study participants.

356 In order to measure knowledge about the application domain (in our case:
 357 complaints handling), we asked participants whether they had previous expe-

358 rience with complaints handling processes (yes/no). Since we did not expect
359 much domain knowledge in a student population, the use of a more exten-
360 sive scale (like the one described in [26]) was not considered. In order to
361 measure respondents' familiarity with the EPC notation, we adapted a three-
362 item scale for notation familiarity developed by Recker [42], which is based
363 on Gemino and Wand's three pre-test questions about the *familiarity*, *compe-*
364 *tence*, and *confidence* of respondents with respect to an analysis method (see
365 Appendix 5.3 and [26]). Accordingly, the three-item familiarity scale assesses
366 familiarity with the (EPC) process modeling notation in a sense of generally
367 felt familiarity (Fam1), self-perceived competence with the notation (Fam2)
368 and self-perceived confidence in using the notation (Fam3). Appendix 5.3 lists
369 all items used in the questionnaire.

370 The second part of the questionnaire shows the process model as depicted in
371 Figure 1. In order to gather data to examine hypothesis **H1**, the participants
372 were asked to identify the top three activity labels that they consider to be the
373 most ambiguous. In the third part, we sought to gather data to examine hy-
374 potheses **H2a**, **H2b** and **H2c**. In order to evaluate usefulness perceptions, we
375 developed a two-item measurement scale that stresses the act of understand-
376 ing. Specifically, we used the Perceived Usefulness (PU) scales developed by
377 Maes and Poels [43] as a basis for our measurement development. The motiva-
_____ *be an improvement to a textual description of the business process*, which essentially
is a tautology. Also note that we focus on perceived usefulness in our experiment
for its importance as a key antecedent to actual usage [44]. The research by Maes
and Poels [43] is much broader in its goal to reveal the contribution of different
dimensions to the quality of conceptual models.

³ We chose not to adapt the PU1 item from [43]. This item cannot be reasonably
applied to text labels. The item would have read *Overall, I think the [label] would*

378 tion is that their PU measures were developed specifically for the conceptual
379 modeling context. Our scales were worded “Overall, I found [label] useful for
380 understanding the process modeled” and “Overall, I think [label] improves my
381 performance when understanding the process modeled”.³ We asked the par-
382 ticipants for their perception in these terms of six activity labels from the
383 process model, using a 7-point Likert scale with the anchor points “Disagree
384 strongly” and “Agree strongly”.

385 We chose not to measure perceived usefulness for each of the twelve distinct
386 labels shown in Figure 1 but instead to record these measures for six la-
387 bels only. We have done so for the pragmatic reason of not making our data
388 collection instrument – and the conduct of the experiment – unnecessarily
389 long. Considering six labels allowed us to obtain 6 (labels) x 29 (number of
390 responses) = 174 data points for hypothesis testing, which we deemed suffi-
391 cient for our analysis. We arbitrarily selected two labels for each of the three
392 styles we identified in the previous section, these being *register receipt date*
393 *of complaint letter* and *inform complainant* as verb-object labels, *registration*
394 and *follow up* that follow the action-noun style, as well as *archiving system*
395 and *incident agenda* for the rest group. We consider our selection strategy
396 sufficiently randomized based on the observation that neither our research ob-
397 jectives nor our hypotheses address the choice of word items or the specificity
398 of the word items used within these labels. Hence, there was no motivation
399 for us to prefer any particular label over another.

400 3.2 Results

401 **Demographics.** The questionnaire of our survey was filled out by 29 students
402 who were at that time following a post-graduate course on process modeling
403 at Eindhoven University of Technology in the Netherlands. Participation was
404 voluntary, and as a reward we offered the students a copy of the study results.
405 25 participants were male, while 4 were female. While some of the participants
406 only had followed university courses for one year, most of them had done so for
407 three years or more, with 3.8 years of study being the mean value. Half of the
408 population had preliminary experience with business process modeling, either
409 professionally or through previous courses. Four persons had not yet worked
410 with EPCs, but the average participant had known them for three months and
411 created 10 models so far. Altogether, 25 out of the 29 participants self-assessed
412 their familiarity with EPCs as better than 3 (average total factor score), with
413 the median being 4.5. We included a brief description of the EPC notation
414 similar to [45, p. 36] such that the participants would in any case have the
415 necessary background to understand the process model. Finally, there were six
416 persons who had some preliminary knowledge of complaint handling processes.

417 Overall, the study population contained individuals with some application do-
418 main knowledge and familiarity of the EPC notation, but without high levels of
419 either. Studies using students have been often criticized for lack of external va-
420 lidity. Despite this criticism, we agree with Gemino and Wand [26,46], Recker
421 and Dreiling [11] as well as Batra et al. [47] that the selection of students
422 over practitioners in this type of research can in fact be advisable. Results
423 from both domain understanding and problem solving tasks could have been
424 confounded by participants that are able to bring to bear prior application

425 domain knowledge in one of the areas [48]. Also, post-graduate students (like
426 the one participating in our study) have been found to be adequate proxies
427 for analysts with low to medium expertise levels [46,49].

428 **Perceived Ambiguity.** The second part of the questionnaire focused on the
429 relationship between label types and perceived ambiguity, as stated in hy-
430 pothesis **H1**. We asked the participants to identify those three activity labels
431 that they consider to be the most ambiguous. Since there are 12 distinct la-
432 bels in the model and 29 participants, we received 348 assessments whether
433 a particular label (belonging to a certain label type) was considered to be
434 among the three most ambiguous ones. The labels *incident agenda*, *complaint*
435 *analysis*, and *archiving system* were mentioned most frequently (14, 13, and
436 12 times). Note that the first and third label belong to the rest group, while
437 *complaint analysis* follows the action-noun style. In contrast, the most am-
438 biguous label following the verb-object style – *inform complainant* – received
439 only two counts overall. The estimated probability of a label for being men-
440 tioned among the three most ambiguous ones was 0.13 for verb-object labels,
441 0.24 for action-noun labels, and 0.45 for the rest group. The 95% confidence
442 intervals show little overlap: 0.08 to 0.19 for verb-object label, 0.17 to 0.31
443 for action-noun labels, and 0.32 to 0.58 for the rest, which correspond to our
444 expectations. To calculate reliability of the assessments made by the study
445 participants, we calculated Cohen’s Kappa [50] statistic to examine the level
446 of agreement between study participants on which labels were most ambigu-
447 ous. The Kappa statistic measures inter-rater reliability whilst controlling for
448 change agreement, and is the generally agreed to be the most adequate tool
449 to measure inter-rater reliability [51]. We obtained a Kappa value of 0.607,
450 which can be classified as substantial or good [51].

451 As per our hypothesis **H1**, we were interested in testing whether the differ-
452 ences between the label types as noted are significant. An analysis of variance
453 (ANOVA) test was not applicable, since the variance of the variable values
454 is not homogeneously distributed and because the dependent variable is not
455 on scale level. Instead, we applied Friedman's two-way analysis of variance
456 by ranks [52]. For each participant, we determined an individual ranking of
457 the three label types. This was achieved as follows. For each label type, we
458 determined its relative proportion among the labels that were rated as most
459 ambiguous by that participant. This gives us 29 matched evaluations, leading
460 to rank totals for the three label types as shown in Table 2. As can be seen,
461 verb-object labels receive the lowest rank total, which means that this type is
462 least often considered as containing ambiguous labels. We advance the null hy-
463 pothesis that there are no differences in individual rankings of the three label
464 types, i.e., that each label type would be mentioned similarly in the top three
465 lists in each of the 29 evaluations. In seeking to refute this null hypothesis,
466 we computed the Friedman statistic χ_r^2 . Note that the Friedman statistic χ_r^2
467 is distributed approximately as chi square [52, p.168]. For this case, it turns
468 out that $\chi_r^2 = 6.28$ with $df = 2$, which means a significant difference in the
469 rankings of the three labeling styles at a 95% confidence level. This result
470 lends support to hypothesis **H1**. We conclude that verb-object style labels are
471 indeed least frequently perceived as being ambiguous, followed by action-noun
472 style labels, and finally rest labels.

473 **Perceived Usefulness.** In the third part of the questionnaire, we recorded
474 the perceived usefulness of six activity labels, two for each label type. We used
475 two measures for PU as described above. More specifically, the used scales
476 measure the extent to which a label is *useful for understanding* and *improves*

| | <i>Verb-Object Labels</i> | <i>Action-Noun Labels</i> | <i>Rest</i> |
|-----------------------|---------------------------|---------------------------|-------------|
| Observed ranked total | 49 | 57 | 68 |
| Expected ranked total | 58 | 58 | 58 |

Table 2

Rank totals for the three label types

477 *the performance when understanding*. We received 174 responses (6×29) that
 478 we were able to link to label types. Based on this data, we examined the
 479 hypotheses **H2a**, **H2b** and **H2c**.

480 Before proceeding with hypothesis testing, we first examined reliability and
 481 validity of the PU measures used. Reliability refers to the internal consistency
 482 of scales. The most widely used test for internal consistency is Cronbach's
 483 α , which should be higher than 0.8 [53]. A second test uses the composite
 484 reliability measure p_c , which represents the proportion of measure variance
 485 attributable to the underlying trait. Scales with p_c values greater than 0.5 are
 486 considered to be reliable [44]. For the PU measures, we obtained a Cronbach's
 487 α value of 0.857, and a p_c value of 0.884, suggesting adequate reliability of the
 488 measures. To establish validity of the measures, we examined convergent and
 489 discriminant validity of the PU measures. Convergent validity can be tested
 490 using three criteria suggested by Fornell and Larcker [54]:

- 491 (1) All indicator factor loadings should be significant and exceed 0.6.
- 492 (2) Construct composite reliabilities p_c should exceed 0.8.
- 493 (3) Average variance extracted (AVE) by each construct should exceed the
 494 variance due to measurement error for that construct (i.e., AVE should
 495 exceed 0.50).

496 Factor loadings for the two PU measures were 0.936 and 0.936 and significant
497 at $p = 0.000$. Composite reliability of the PU construct was estimated to be
498 0.884, and average variance extracted was computed to be 0.936. These re-
499 sults suggest adequate convergent validity. To check for discriminant validity,
500 we considered whether measures used for the PU construct would cross-load
501 on other constructs considered (in our case, measures for notation familiarity).
502 The test for discriminant validity is met when the AVE for each construct ex-
503 ceeds the squared correlation between that and any other construct considered
504 in the factor correlation matrix. The squared correlation between the PU and
505 the familiarity factor were computed to be 0.030, which shows that the AVE
506 measures for both PU (0.936) and notation familiarity (0.927) well exceeded
507 the squared correlation between the factors. Appendix 5.3 summarizes factor
508 loadings, communalities, and correlations.

509 Next, to test the hypotheses, we first constructed a box-plot for the average
510 total factor scores for the PU variable, and examined the rank correlations as
511 well as the differences in variance between the average total factor scores for
512 the different label types. Figure 2 gives the box plots.

513 As illustrated by the box-plot in Figure 2, verb-object labels were found to
514 be best in terms of their perceived usefulness, followed by action-noun labels,
515 and then the rest group. Perusal of Table 3 further shows that the reported
516 95% confidence intervals around the means hardly overlap between the label
517 types. In particular, the verb-object style can easily be distinguished from
518 the action-noun style: The upper bounds of the confidence intervals for the
519 action-noun style are strictly lower than the lower bounds for the verb-object
520 style. These results lend initial support to hypotheses **H2a**, **H2b** and **H2c**.

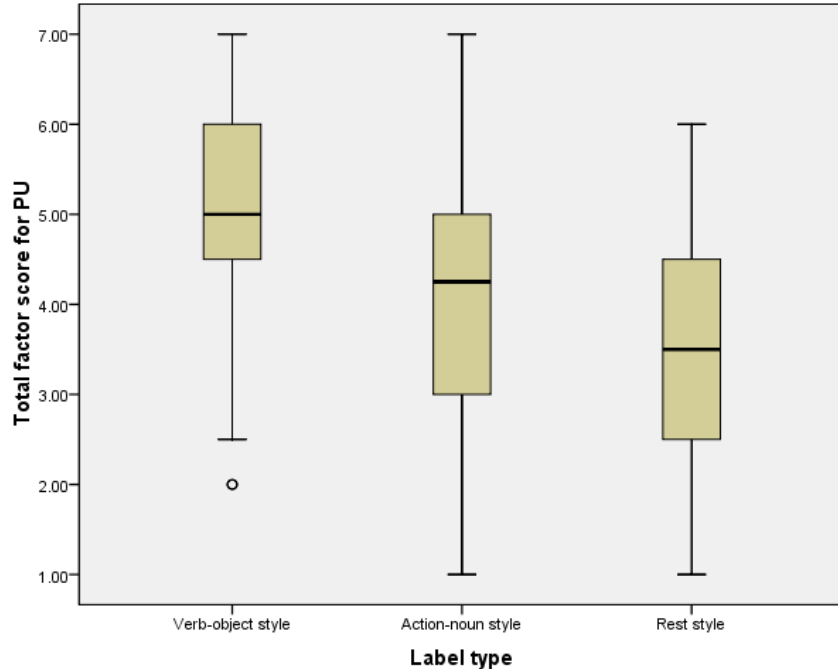


Figure 2. Box-plot of perceived usefulness rankings, by label type

521 As a next step, we examined whether the noted differences are statistically
 522 significant. In the data, we identified a significant negative Spearman rank
 523 correlation between the label style and its perceived usefulness (-0.430 at
 524 99% significance level). This finding suggests that a deviation from the verb-
 525 object style to any of the other two is connected with lower usefulness per-
 526 ceptions, hence lending further support to hypotheses **H2a**, **H2b** and **H2c**.
 527 Additionally, based on the data displayed in Table 3 we performed an analysis
 528 of variance (ANOVA) test implemented in SPSS 16.0 [55] to further exam-
 529 ine the differences in the average total factors scores for PU. Between-group
 530 differences across the different label styles were statistically significant with
 531 $F = 18.495, p = 0.000$, thereby confirming our test results.

532 To test whether there are significant pair-wise differences between the label
 533 types (verb-object versus action-noun, verb-object versus rest, and action-

| | | Perceived Usefulness (avg. total factor score) |
|--------------------|-----------------|--|
| <i>Verb-object</i> | 95% upper bound | 5.304 |
| | Mean | 5.000 |
| | 95% lower bound | 4.696 |
| <i>Action-noun</i> | 95% upper bound | 4.480 |
| | Mean | 4.121 |
| | 95% lower bound | 3.761 |
| <i>Rest</i> | 95% upper bound | 3.905 |
| | Mean | 3.552 |
| | 95% lower bound | 3.199 |

Table 3

Perceived Usefulness of Label Types

534 noun versus rest), we repeated the ANOVA analysis using the Contrast func-
535 tion [55] to detect pair-wise differences. For perceived usefulness, the contrast
536 between verb-object and action-noun style was significant at $contrastValue =$
537 $0.879, t = 3.665, p = 0.000$, while the contrast between verb-object and rest
538 style was significant $contrastValue = 1.448, t = 6.036, p = 0.000$. Finally, the
539 contrast between action-noun and rest style was significant at $contrastValue =$
540 $0.569, t = 2.371, p = 0.019$. These results further lend strong support to hy-
541 potheses **H2a**, **H2b** and **H2c**. In summation, the reported findings support
542 our hypotheses **H2a**, **H2b** and **H2c** that verb-object styles are regarded more
543 useful than action-noun styles, and rest styles.

544 **Perceived Ambiguity's effect on Perceived Usefulness.** As discussed in
 545 the hypothesis development section, our study rests on the assumption that
 546 ambiguity of textual labels is an impediment to the perceived usefulness of
 547 the label for understanding the process modeled. To test this assumption as
 548 specified in hypothesis **H3**, we once again performed an ANOVA test.

549 Support for hypothesis **H3** exists if there are significant differences in the
 550 average total factor scores for perceived usefulness for labels that are either
 551 considered ambiguous, or not, with the expectation that the average total
 552 factor score will be lower for the group that considered a particular labeling
 553 style to be ambiguous. Prior to conduct, ANOVA assumptions were tested
 554 and showed no violation. Table 4 provides the results.

| | Unambiguous label <i>N=132</i> | | Ambiguous label <i>N=42</i> | | ANOVA | |
|-------------------------|--------------------------------------|-------|-----------------------------------|-------|--------|-------|
| | Mean | StDev | Mean | StDev | F | Sig. |
| Perceived usefulness | 4.538 | 1.241 | 3.238 | 1.495 | 31.553 | 0.000 |

Table 4

Average Perceived Usefulness scores for Ambiguous versus Unambiguous Label
Types

555 The results displayed in Table 4 confirm our assumption and lend strong sup-
 556 port to hypothesis **H3**. The average total factor score for perceived usefulness
 557 was higher for those label types that were not listed as ambiguous by the par-
 558 ticipants (reported average total factor scores are 4.538 in contrast to 3.238).

559 The ANOVA test showed these differences to be statistically significant at
560 $p = 0.000$.

561 **Moderating Effects.** As discussed in the demographics section, the partici-
562 pants ranged in terms of their familiarity with the EPC notation used in the
563 process model, as well as in their knowledge of the chosen application domain
564 (complaints handling). More precisely, six participants brought to bear experi-
565 ence with complaints handling domain, and 17 out of 29 participants were
566 above the median in notation familiarity.

567 Again we first established reliability and validity of the measure “familiarity
568 with the EPC notation”. Cronbach’s α for the familiarity scale was computed
569 to be 0.914, and composite reliability was computed to be 0.859. Factor load-
570 ings for the three familiarity measures were 0.919, 0.930 and 0.931, all sig-
571 nificant at $p = 0.000$. Average variance extracted of the familiarity construct
572 was estimated to be 0.927. As described above, AVE also exceeded the squared
573 correlation between the PU and the familiarity construct. Altogether, these re-
574 sult suggest adequate reliability and validity. Appendix 5.3 summarizes factor
575 loadings, communalities, and correlations.

576 In order to test hypotheses **H4a** and **H4b**, we examined the differences in the
577 average total factor scores for perceived usefulness of the labels between two
578 sets of two groups of participants (high/low application domain knowledge
579 and high/low familiarity with the EPC notation). Support for the hypotheses
580 would then exist if the differences in the dependent variables between the
581 groups would be significant. We used an analysis of covariance (ANCOVA) test
582 implemented in SPSS 16.0 to test the hypotheses. ANCOVA is an appropriate
583 analysis technique because it allows to control for potential effects of covariates

584 in the examination of dependent variable scores between two treatment groups
585 [55]. ANCOVA assumptions of equal slopes were tested prior to conduct, and
586 showed no violation of normality.

587 We used two covariates in the analysis of the effect on labeling type on per-
588 ceived usefulness. The first is the binary variable “Knowledge of the com-
589 plaints handling domain”, which simply establishes the existence of any rel-
590 evant knowledge in this domain. As a second covariate, we used the median
591 of the total factor score of the three item “Familiarity” scale, to separate the
592 respondents pool in two groups using a dummy variable (high familiarity/low
593 familiarity). Both variables have been described in Section 3.1. Appendix 5.3
594 lists all items used in the questionnaire. We obtained the following results:

- 595 - Application domain knowledge does not show a significant interaction ef-
596 fect on the relationship between label type and perceived usefulness ($F =$
597 $1.363, p = 0.245, \text{partialetasquare} = 0.008$). Accordingly, hypothesis **H4a**
598 must be refuted.
- 599 - Notation familiarity does not show a significant interaction effect on the
600 relationship between label type and perceived usefulness ($F = 1.334, p =$
601 $0.239, \text{partialetasquare} = 0.006$). Accordingly, hypothesis **H4b** must be
602 refuted.

603 These results are similar to those reported in [11,26], which also did not in-
604 dicate significant moderation effects of their measures of application domain
605 knowledge or familiarity with the notation on understanding of conceptual
606 models – and contrary to those reported in [38,39], both of which reported
607 some spurious effects on a number of the dependent variables they consid-
608 ered. In the context of the study reported in this paper, the results indicate

609 that understanding of textual labels contained in process models is indepen-
610 dent from any expertise gained from previous notation usage or from previous
611 knowledge of the considered domain. In light of the other results presented
612 above, the findings suggest that a label’s usefulness is indeed dependent on
613 the grammatical style of the labels itself.

614 3.3 Discussion

615 The support for our hypotheses strongly suggests that a verb-object labeling
616 style is rightfully proposed as a preferred way of activity labeling. Indeed,
617 our results indicate strong and favorable perceptions towards a superiority
618 of the verb-object labeling style. Given the key role that usage beliefs (such
619 as perceived ambiguity or perceived usefulness) play in informing actual us-
620 age behavior [44,56,57], we deem this finding instrumental to explaining, and
621 supporting, process model understandability. However, whilst process model-
622 ers tend to favor verb-object styles, this situation does not necessarily reflect
623 actual usage for activity labeling. In fact, our exploration of the usage fre-
624 quency of activity labels in the SAP reference model indicates that a large
625 proportion of labels found in practice *cannot* be interpreted as genuine im-
626 plementations of this style (see Section 2). In contrast, our results indicate
627 that there is wide variety in labeling. We would argue that this situation can
628 largely be attributed to a lack of operationalized guidance in the proper use
629 of “good” labeling styles (such as the verb-object style). We further argue
630 that the results from our empirical investigations have implications for both
631 research and practice on the quest towards guiding process modelers towards
632 the consequent and consistent usage of labeling styles. In the next section, we

633 will address this issue in more detail.

634 4 Implications

635 In this section we highlight some implications of this research. We first discuss
636 implications for research in Section 4.1 and then implications for practice in
637 Section 4.2.

638 4.1 For Research

639 Our research has strong implications for research into process modeling. While
640 there are some works that describe process models and the information they re-
641 veal in a holistic way, e.g. [16], most contributions consider a process model as
642 a structural design artifact. This holds for the whole stream of formal analysis
643 techniques, such as those based on Petri nets. The latter stream has domi-
644 nated, for example, papers presented at the recent Business Process Manage-
645 ment conferences (see [58,59]). In that type of work, textual labels are usually
646 little more than identifiers to the activity concepts in a process model. Our
647 research, in contrast, shows the relevance of labeling for perceived ambiguity
648 and perceived usefulness – that is, to key beliefs informing actual usage behav-
649 ior. While this does not directly allow a statement on the relative importance
650 of structure and labeling for the pragmatic value of a process model, it sug-
651 gests that both aspects are complementary and hence deserve closer attention
652 in process modeling research.

653 Furthermore, in our research we designed a measurement for perceived am-
654 biguity for the textual content of activity labels. It is worth noting that this

655 concept can be equally adapted for investigating the structural elements of a
656 process model. For instance, the formal semantics of OR-join elements have
657 been debated extensively in process modeling research (see e.g. [60,61] for an
658 overview of the discussion). The problem with the OR-join is that it is meant
659 to synchronize only those branches that are still active – which may lead to
660 contradictions when multiple OR-joins wait for each other. Up until now, no
661 generally accepted formalization has been found for this element. It would be
662 an interesting matter for future research to investigate in how far this formal
663 ambiguity materializes in user perceptions on ambiguity.

664 We see another research implication stemming from the fact that the use of
665 textual labels in conceptual models addresses aspects of linguistics. Through
666 our grammatical analysis we were able to show how some linguistic principles
667 inform information systems practice in the conceptual modeling space. This
668 situation suggests the field of linguistics to be a fruitful reference discipline
669 from which theories, research strategies as well as empirical measurements
670 could be drawn that advance our understanding of the role and practice of
671 conceptual modeling. In our work, we relied in part on existing measurements
672 adopted from information systems research to measure usefulness and ambi-
673 guity of textual labels. Clearly, these types of evaluations also pertain to the
674 study of language in general. Future research could examine to what extent
675 knowledge advances from the field of linguistics could inform practices and out-
676 comes in information systems and conceptual modeling. And indeed, a number
677 of scholars have already established that linguistics contributes to informing
678 the body of knowledge around conceptual modeling, e.g., [62,63,64,65].

679 4.2 For Practice

680 In addition to this work's academic merits we also identify a number of impli-
681 cations that pertain directly to process modeling practice. Most notably, our
682 research confirms the suggestion – hitherto largely unreflected – that verb-
683 object styles are an appropriate labeling convention.

684 In order to lend better support for practitioners working with verb-object la-
685 beling styles, it is important to remember that the labeling of graphical model
686 constructs refers to the act of annotating the model construct with information
687 about the intended real-world *domain semantics* that the graphical element is
688 purported to articulate in the model. Domain semantics define the real-world
689 meaning, or essence thereof, of the terms used in any conceptual model, that
690 is, of words and phrases used to label constructs [17]. The delicate part is
691 that some of these semantics are well-known and unambiguous while others
692 may vary with context, i.e., they can be subject to multiple interpretations.
693 Furthermore, the words used to annotate semantics (e.g., the verb and the
694 object in the verb-object labeling style), are typically selected by the model
695 developer, sometimes without a great deal of thought spent on finding the one
696 that best reflects the intended real-world semantics. This can lead to prob-
697 lems when reading and interpreting the model, especially – as often found
698 in modeling practice – model developer and reader audiences do not overlap.
699 This situation is further complicated by the vast amount of terms found in a
700 natural language such as English. For example, the online catalogue WordNet
701 contains over 21,000 different verb word forms alone [35].

702 In essence, our research results imply, and highlight, a need for a closer inte-

703 gration of process model and structured content. This would have to be sup-
704 ported by process modeling tools. We now discuss this issue in more detail.
705 In particular, in Section 4.2.1 we analyze *label parsing* as a quality measure
706 mechanism to enforce the verb-object style. Sections 4.2.2 and 4.2.3 investi-
707 gate how controlled vocabularies for *business activities* and *business objects*
708 – the two central terms in the verb-object style – can be supported in pro-
709 cess modeling. The key challenge in this area is to aid the process modeler in
710 determining a precise label. In linguistic terms, this challenge closely relates
711 to the problem of synonyms and homonyms. In order to avoid interpretation
712 problems syntactically different terms should be used for referring to seman-
713 tically distinct verbs or object (avoid synonyms) and syntactically equivalent
714 terms should represent equivalent semantics (avoid homonyms).

715 4.2.1 *Parsing of labels*

716 Current approaches to establishing syntactical correctness of process models
717 (e.g., [66,67,68]) only consider properties of the process graph. The results of
718 our empirical research suggest extending these approaches to also consider the
719 labeling style of activities. This requires a grammatical analysis of the activity
720 labels. The respective parsing can be facilitated in two levels of sophistication.
721 In the simple case, the different words of the text label string are identified
722 using standard programming facilities such as the Java String Tokenizer. Tools
723 like WordNet [35] then check whether the first word is a verb or not, and
724 whether some of the remaining words are nouns. A business process modeling
725 tool can use such an analysis for pointing modelers to labels that do not
726 follow the verb-object style. This approach can be extended by using verb
727 phrase parsing techniques such as [69] to identify the grammatical role of each

728 word in the text label. This way, a process modeling tool can help the user
729 not only to use verb-object style, but also to avoid grammatical errors.

730 The enforcement of the verb-object style in this way might help to close the
731 gap between natural language and formal language processing. And indeed, the
732 relationship between process models and natural language has been discussed
733 and utilized in various works. In [70] the authors investigate in how far the
734 three steps of building a conceptual model (linguistic analysis, component
735 mapping, and schema construction) can be automated using a model for pre-
736 design. Further text analysis approaches have been used to link activities in
737 process models to document fragments [71] and to compare process models
738 from a semantic perspective [72]. Most beneficiary is the verb-object style for
739 model verbalization and paraphrasing, see [73,74]. Such verbalization is an
740 important step in model and requirements validation [75]. For instance, verb-
741 object style labels can easily be verbalized using the “You have to” prefix
742 which yields natural language sentences like “You have to reject order”. This
743 way, verb-object labeling and automatic parsing enables a better validation of
744 process models.

745 4.2.2 *Controlled Object Vocabulary*

746 The parsing of activity labels can then be used to introduce further measures
747 of quality assurance. In this context, it is important that the entities referenced
748 in the labels of the process model relate to relevant concepts of the organi-
749 zation and its environmental context. In research and practice it is widely
750 acknowledged that an agreed set of key terms is an essential prerequisite for
751 modeling business processes [76]. The existence of a repository of such terms

752 and entities, and its integration of the process modeling tool is essential for
753 supporting respective quality measures. In the following paragraph, we discuss
754 two options for integrating such a repository: by modeling and by reuse.

755 Different authors, e.g., [76], recommend a preparatory step called technical
756 term modeling before modeling the actual process. Technical term models
757 capture the key entities involved in a business process and delineate their
758 hierarchy and semantic relationships. Often, entity-relationship diagrams or
759 class diagrams are employed as a modeling language by practitioners. Some
760 niche tools like Semtalk already support an integration of ontologies and pro-
761 cess models for this purpose [77], but major tools such as ARIS or Telelogic
762 System Architect do not.

763 Instead of modeling from scratch there is also the option to reuse existing data
764 dictionaries and data models. These do not necessarily have to be company-
765 specific. Domain standards and ontologies like the XML Common Business
766 Library (xCBL), or Health Level 7 (HL7) are well suited. Also, some authors,
767 e.g., [78] have suggested general ontological models to guide the act of con-
768 ceptual modeling by defining key terms and concepts and their relationships.
769 In addition to this work, tools such as the WordNet catalogue can be used
770 to resolve homonym and synonym conflicts between the users data input for
771 the label and the data dictionary. Furthermore, a side effect of this concept
772 could be a better integration of process modeling and semantic business pro-
773 cess management (SBPM) technologies [79]. And indeed, previous research
774 has already delivered stimulating work towards a better integration between
775 organizational objects and data concepts, and their role in dynamic processes.
776 Wagner [80], for instance, describes how static, dynamic and deontic aspects
777 of organizations can conceptually be captured on the basis of a set of 19 on-

778 tological principles. It is on the basis of work like this where future research
779 can deliver relevant guidance to process modeling practice.

780 4.2.3 *Controlled Verb Vocabulary*

781 While class diagrams and data dictionaries can be easily used to control the
782 object part of verb-object labels, the case is more difficult for verbs. Some
783 work in data modeling has shown that the variety in relationship semantics is
784 much smaller than the potential set of relationship labels [17]. Consequently,
785 Storey argues for, and develops, an ontology for the semantic classification of
786 relationship-type constructs in data models based on dictionaries and business
787 taxonomies. We argue that a similar idea is applicable for activity labeling in
788 process modeling.

789 Some work is available as a foundation for such an endeavor. The MIT process
790 handbook [20], for instance, discusses a wide range of action terms to be
791 used in business contexts. Building on the lexical database WordNet [35],
792 the MIT handbook defines an inheritance hierarchy that originates from eight
793 generic verbs (*viz.*, *create*, *modify*, *preserve*, *destroy*, *combine*, *separate*, *decide*,
794 *manage*). Verb classifications and verb ontologies have been proposed before.
795 The systematic work by Levin is an important contribution in this area. It
796 defines 49 semantic classes of verbs and categorizes more than 3,000 English
797 verbs [81]. A formal approach towards a verb ontology is reported in [82].
798 Yet, there are several problems, in particular with the classifications of the
799 MIT process handbook and that by Levin when applied to process modeling.
800 Recent research has shown that both schemes cover only a limited amount
801 of verbs found in real-world process model collections (44% and 68%) [83].

802 Furthermore, neither of the hierarchies is a tree, a problem that stems from
803 synonyms and homonyms.

804 Clearly, future work is required to address these shortcomings. More precisely,
805 to lend further support to labeling practices, a verb hierarchy should be con-
806 structed that defines generic verb terms of pertinence to business process con-
807 texts whilst avoiding homonyms and synonyms at all. Such a verb hierarchy
808 could then become an integral part of process modeling tools, as much as the-
809 sauri are used in word processing tools. Facilities to extend this hierarchy with
810 domain-specific verbs could be implemented in as simple a way as defining user
811 extensions to the general verb dictionary.

812 **5 Conclusions**

813 This section concludes the paper by summarizing the contributions, the limi-
814 tations, and by giving an outlook on future research.

815 *5.1 Contributions*

816 In this paper we discussed an essential yet under-researched aspect of process
817 modeling practice, namely that of labeling the graphical activity elements in
818 a process model. In this way, we complement the existing streams of research
819 investigating other dimensions of process modeling (e.g., the data, resource,
820 or control-flow perspectives). Our line of research is based on the assumption
821 that process model understanding can be improved if a more systematic way
822 of labeling constructs can be found. Based on Dual Coding Theory and CTML
823 we argued that understanding can be improved if more consideration is given

824 to the *style* and *choice of terms* for labeling activities in process models. We
825 examined over 600 process models and considered data on the user percep-
826 tions of labeling styles to lend support for our arguments. We then explored
827 the implications of our empirical findings and suggested three programs of re-
828 search towards better, and more stringent, support for process model labeling
829 practices.

830 5.2 Limitations

831 Clearly, our research has its limitations and is not yet complete. First, even
832 though we examined a considerably large number of process models, we only
833 considered EPCs of the SAP reference model. This may limit the extent to
834 which our results can be generalized. However, the way we described the de-
835 sign of our experiments, and the inclusion of our data collection instrument
836 (see Appendix 5.3), will allow researchers to replicate our study in other pro-
837 cess modeling contexts, e.g., using different model sets, or different process
838 modeling notations such as BPMN or UML Activity Diagrams. It will be an
839 interesting topic of future research to examine whether our findings can be
840 directly transferred to other activity-based process modeling languages, for
841 instance, those that do not explicitly label events (e.g., BPMN, YAWL).

842 Second, in our study we examined the general labeling style used in process
843 modeling. Clearly, not only the style of labeling but also the specificity of
844 the word items used within these styles (e.g., the actual verb or object terms
845 used in the verb-object style) will have an influence on the quality of the
846 model produced. Future research should thus more closely investigate how
847 the choice or specificity of terms influence process model quality. We outlined

848 some suggestions for such research in Section 4.1.

849 Third, as with any other research studying perceptual beliefs, our measure-
850 ment strategy is a potential source of limitations. For some of the aspects (e.g.,
851 ambiguity) we considered in our study, we had to develop new measures. In
852 the case of perceived ambiguity, the chosen operationalization only allowed us
853 to calculate Cohen's Kappa as a measure of reliability. In effect, we cannot
854 rule out potential validity issues with this measure. For other aspects (e.g.,
855 perceived usefulness, familiarity), in part, we took inspiration from existing
856 measures used in other studies of Information Systems or conceptual mod-
857 eling phenomena, to propose modified measurement items that we deemed
858 appropriate to our problem context. While reliance on existing measurements
859 in instrument development may be an efficient research practice, it should
860 not be considered superior to rigorously developing new measures [84]. For
861 instance, an alternative to our approach may have been the development of
862 a specialized scale for measuring the understandability of labels, which may
863 have given insights beyond the ones presented in this paper. Furthermore, our
864 approach to *modify* existing measures to make them fit to our research context
865 clearly hampers reliance on their earlier validations. For this reason, we have
866 given careful consideration to testing potential reliability and validity issues
867 (see the analyses reported in Appendix 5.3). While our test results indicate the
868 adequateness of the selected operationalizations, we can imagine that further
869 studies will be useful for a proper reflection on our measurement strategy, and
870 we would like to invite our fellow scholars to join in this endeavor.

871 5.3 Outlook

872 Some of the future research streams we consider will be as follows. Aside from
873 seeking to validate our findings on a more general level by considering vari-
874 ous other process modeling notaitons in use today (e.g., BPMN, YAWL), we
875 also aim to examine empirically the suitability of different verb classification
876 schemes for classifying activity tasks in process models. Similar to the exper-
877 iment described in [17], we will have respondents classify activity tasks in a
878 number of process models as per the verb classification schemes to establish
879 the viability of these schemes.

880 In a related stream of research we will then aim to establish in an empir-
881 ical setting whether the consistent usage of the operationalized verb-object
882 style in process models does *in fact* warrant improved model understandabil-
883 ity. CTML suggests three outcomes of understanding – retention, recall and
884 transfer – that can be used as measures in a related empirical study. In con-
885 ducting such a study we can refer to the works of Gemino and Wand [26]
886 and Recker and Dreiling [11] that both used these measures for examining
887 understanding generated through data [26] and process modeling [11], respec-
888 tively. The empirical results reported in this paper show that label styles have
889 an effect on user perceptions of usefulness, and in our future work we are
890 keen to examine the effect of labeling styles on actual measures of usability,
891 understanding and performance.

892 Finally, our research into the labeling of graphical elements should lead to spe-
893 cific guidelines that can be effectively used by modelers. Even if a verb-object
894 style of modeling is preferable over other styles, clearly more perspectives,

895 e.g. data, resource, and control-flow, should be considered to create an overall
896 understandable model. Earlier research, for example, has shown the impact of
897 a process model's size, structure, and modularity on its overall understand-
898 ability [9,85]. Based on such insights, a preliminary, broad set of guidelines
899 is presented in [86]. This so-called 7PMG set includes a guideline on using
900 the verb-object style, as well as guidelines on the number of elements in a
901 model, the application of structured modeling, and the decomposition of a
902 process model. Aside from the challenge to generate guidelines from emerging
903 research on process model quality, a whole new venue of research emerges with
904 respect to establishing the effectiveness of such a set. A potential source of
905 inspiration is the field of use case writing, which we referred to earlier (see
906 Section 2.2). In various papers in this field, experiments are described to as-
907 sess the impact of modelers' usage of guidelines on the quality of use case
908 descriptions and how alternative guideline sets compare with each other in
909 this respect [27,28,87].

910 **Acknowledgement**

911 The authors thank the anonymous reviewers for their valuable feedback, which
912 improved the paper significantly.

913 **Appendix I: Questionnaire Material Used**

914 **Demographics**

- 915 • Gender (male/female)
- 916 • Years of tertiary education
- 917 • Working experience in process modeling (yes/no)

- 918 • Experience in EPC modeling (months)
- 919 • Number of EPC models created/read
- 920 • Training received in EPCs (formal/internal/university/on the job/auto-didact/reading/other)

921 **Familiarity with the EPC notation (7-point scale from “Strongly**
922 **disagree” to “Strongly agree”)**

- 923 • Overall, I am very familiar with EPCs.
- 924 • I feel very confident in understanding process models created with EPCs.
- 925 • I feel very competent in using EPCs for process modeling.

926 **Application domain knowledge**

- 927 • Knowledge of claims handling processes (yes/no)

928 **Perceived Ambiguity of a Label**

- 929 • Please list the three function labels from the model that you consider to be
930 the most *ambiguous* ones, i.e., they are most open for alternative inter-
931 pretations:

932 **Perceived Usefulness of a Label (7-point scale from “Strongly dis-**
933 **agree” to “Strongly agree”)**

- 934 • Overall, I found Label *X* useful for understanding the process modeled.
- 935 • Overall, I think Label *X* improves my performance when understanding the
936 process modeled.

937 Appendix II: Reliability and Validity Results

| | Perceived usefulness | | Notation familiarity |
|----------------------------------|-------------------------|----------------------|---|
| | <i>Factor loadings</i> | <i>Communalities</i> | <i>Factor loadings</i> <i>Communalities</i> |
| PU1 | 0.936 | 0.877 | |
| PU2 | 0.936 | 0.877 | |
| Fam1 | | | 0.919 0.845 |
| Fam2 | | | 0.930 0.866 |
| Fam3 | | | 0.931 0.867 |
| Cronbach's Alpha | 0.857 | | 0.914 |
| Composite reliability | 0.884 | | 0.868 |
| Average Variance Extracted | 0.936 | | 0.927 |
| Correlation | 1.000 -0.030 | | 0.030 1.000 |

938 **References**

- 939 [1] I. Davies, P. Green, M. Rosemann, M. Indulska, S. Gallo, How do practitioners
940 use conceptual modeling in practice?, *Data & Knowledge Engineering* 58 (3)
941 (2006) 358–380.
- 942 [2] M. Dumas, W. M. P. van der Aalst, A. H. M. ter Hofstede (Eds.), *Process
943 Aware Information Systems: Bridging People and Software Through Process
944 Technology*, John Wiley & Sons, Hoboken, New Jersey, 2005.
- 945 [3] T. Erl, *Service-oriented Architecture: Concepts, Technology, and Design*,
946 Prentice Hall, Upple Saddle Revier, New Jersey, 2005.
- 947 [4] C. Ferris, What are web services?, *Communications of the ACM* 46 (6) (2003)
948 31–32.
- 949 [5] B. Curtis, M. I. Kellner, J. Over, Process modeling, *Communications of the
950 ACM* 35 (9) (1992) 75–90.
- 951 [6] A.-W. Scheer, *ARIS - Business Process Modeling*, 3rd Edition, Springer, Berlin,
952 Germany, 2000.
- 953 [7] J. Dehnert, W. M. P. van der Aalst, Bridging the gap between business models
954 and workflow specifications, *International Journal of Cooperative Information
955 Systems* 13 (3) (2004) 289–332.
- 956 [8] D. L. Moody, Theoretical and practical issues in evaluating the quality of
957 conceptual models: Current state and future directions, *Data & Knowledge
958 Engineering* 15 (3) (2005) 243–276.
- 959 [9] J. Mendling, H. A. Reijers, J. Cardoso, What makes process models
960 understandable?, in: G. Alonso, P. Dadam, M. Rosemann (Eds.), *Business
961 Process Management - BPM 2007*, Vol. 4714 of *Lecture Notes in Computer
962 Science*, Springer, Brisbane, Australia, 2007, pp. 48–63.

- 963 [10] J. Mendling, G. Neumann, W. M. P. van der Aalst, Understanding the
964 occurrence of errors in process models based on metrics, in: R. Meersman,
965 Z. Tari (Eds.), *On the Move to Meaningful Internet Systems 2007*, Vol. 4803 of
966 *Lecture Notes in Computer Science*, Springer, Vilamoura, Portugal, 2007, pp.
967 113–130.
- 968 [11] J. Recker, A. Dreiling, Does it matter which process modelling language
969 we teach or use? an experimental study on understanding process modelling
970 languages without formal education, in: M. Toleman, A. Cater-Steel, D. Roberts
971 (Eds.), *18th Australasian Conference on Information Systems*, The University
972 of Southern Queensland, Toowoomba, Australia, 2007, pp. 356–366.
- 973 [12] M. Rosemann, J. Recker, M. Indulska, P. Green, A study of the evolution of
974 the representational capabilities of process modeling grammars, in: E. Dubois,
975 K. Pohl (Eds.), *Advanced Information Systems Engineering - CAiSE 2006*, Vol.
976 4001 of *Lecture Notes in Computer Science*, Springer, Luxembourg, Grand-
977 Duchy of Luxembourg, 2006, pp. 447–461.
- 978 [13] W. M. P. van der Aalst, A. H. M. ter Hofstede, B. Kiepuszewski, A. P. Barros,
979 *Workflow patterns*, *Distributed and Parallel Databases* 14 (1) (2003) 5–51.
- 980 [14] N. Russell, W. M. P. van der Aalst, A. H. M. ter Hofstede, D. Edmond,
981 *Workflow resource patterns: Identification, representation and tool support*,
982 in: O. Pastor, J. Falcao e Cunha (Eds.), *Advanced Information Systems*
983 *Engineering - CAiSE 2005*, Vol. 3520 of *Lecture Notes in Computer Science*,
984 Springer, Porto, Portugal, 2005, pp. 216–232.
- 985 [15] N. Russell, A. H. M. ter Hofstede, D. Edmond, W. M. P. van der Aalst, *Workflow*
986 *data patterns: Identification, representation and tool support*, in: L. M. L.
987 Delcambre, C. Kop, H. C. Mayr, J. Mylopoulos, O. Pastor (Eds.), *Conceptual*
988 *Modeling - ER 2005*, Vol. 3716 of *Lecture Notes in Computer Science*, Springer,
989 Klagenfurt, Austria, 2005, pp. 353–368.

- 990 [16] J. Krogstie, G. Sindre, H. D. Jorgensen, Process models representing knowledge
991 for action: a revised quality framework, *European Journal of Information*
992 *Systems* 15 (1) (2006) 91–102.
- 993 [17] V. C. Storey, Comparing relationships in conceptual modeling: Mapping
994 to semantic classifications, *IEEE Transactions on Knowledge and Data*
995 *Engineering* 17 (11) (2005) 1478–1489.
- 996 [18] L. D. Miles, *Techniques of Value Analysis and Engineering*, McGraw-Hill, New
997 York, New York, 1961.
- 998 [19] A. Sharp, P. McDermott, *Workflow Modeling: Tools for Process Improvement*
999 *and Application Development*, Artech House, Boston, Massachusetts, 2001.
- 1000 [20] T. W. Malone, K. Crowston, G. A. Herman, *Organizing Business Knowledge:*
1001 *The MIT Process Handbook*, The MIT Press, Cambridge, Massachusetts, 2003.
- 1002 [21] A. Paivio, Dual coding theory: Retrospect and current status, *Canadian Journal*
1003 *of Psychology* 45 (3) (1991) 255–287.
- 1004 [22] L. R. Brooks, The suppression of visualization by reading, *The Quarterly*
1005 *Journal of Experimental Psychology* 19 (4) (1967) 289–299.
- 1006 [23] R. E. Mayer, Models for understanding, *Review of Educational Research* 59 (1)
1007 (1989) 43–64.
- 1008 [24] R. E. Mayer, *Multimedia Learning*, Cambridge University Press, Cambridge,
1009 Massachusetts, 2001.
- 1010 [25] F. Bodart, A. Patel, M. Sim, R. Weber, Should optional properties be used in
1011 conceptual modelling? a theory and three empirical tests, *Information Systems*
1012 *Research* 12 (4) (2001) 384–405.
- 1013 [26] A. Gemino, Y. Wand, Complexity and clarity in conceptual modeling:
1014 Comparison of mandatory and optional properties, *Data & Knowledge*

- 1015 Engineering 55 (3) (2005) 301–326.
- 1016 [27] K. Cox, K. T. Phalp, Replicating the crews use case authoring guidelines
1017 experiment, *Empirical Software Engineering* 5 (3) (2000) 245–267.
- 1018 [28] K. T. Phalp, J. Vincent, K. Cox, Improving the quality of use case descriptions:
1019 Empirical assessment of writing guidelines, *Software Quality Journal* 15 (4)
1020 (2007) 383–399.
- 1021 [29] R. B. Davis, *Business Process Modelling with ARIS: A Practical Guide*,
1022 Springer, London, England, 2001.
- 1023 [30] J. Mendling, M. Strembeck, Influence factors of understanding business process
1024 models, in: W. Abramowicz, D. Fensel (Eds.), *Business Information Systems -*
1025 *BIS 2008*, Vol. 7 of *Lecture Notes in Business Information Processing*, Springer,
1026 Innsbruck, Austria, 2008, pp. 142–153.
- 1027 [31] G. Keller, T. Teufel, *SAP R/3 Process Oriented Implementation: Iterative*
1028 *Process Prototyping*, Addison-Wesley, Boston, Massachusetts, 1998.
- 1029 [32] M. Daneva, *Erp requirements engineering practice: Lessons learned*, *IEEE*
1030 *Software* 21 (2) (2004) 26–33.
- 1031 [33] J. Mendling, H. M. V. Verbeek, B. F. van Dongen, W. M. P. van der Aalst,
1032 G. Neumann, Detection and prediction of errors in epcs of the sap reference
1033 model, *Data & Knowledge Engineering* 64 (1) (2008) 312–329.
- 1034 [34] A. Dreiling, M. Rosemann, W. M. P. van der Aalst, W. Sadiq, From conceptual
1035 process models to running systems: A holistic approach for the configuration of
1036 enterprise system processes, *Decision Support Systems* 45 (2) (2008) 189–207.
- 1037 [35] G. A. Miller, *Wordnet - a lexical database for english*, *Communications of the*
1038 *ACM* 38 (11) (1995) 39–41.

- 1039 [36] R. M. W. Dixon, Deriving verbs in english, *Language Sciences* 30 (1) (2008)
1040 31–51.
- 1041 [37] R. Price, G. Shanks, A semiotic information quality framework: Development
1042 and comparative analysis, *Journal of Information Technology* 20 (2) (2005) 88–
1043 102.
- 1044 [38] V. Khatri, I. Vessey, P. C. V. Ramesh, P. Sung-Jin, Understanding conceptual
1045 schemas: Exploring the role of application and is domain knowledge,
1046 *Information Systems Research* 17 (1) (2006) 81–99.
- 1047 [39] K. Masri, D. C. Parker, A. Gemino, Using iconic graphics in entity-relationship
1048 diagrams: The impact on understanding, *Journal of Database Management*
1049 19 (3) (2008) 22–41.
- 1050 [40] D. Batra, J. G. Davis, Conceptual data modelling in database design:
1051 Similarities and differences between expert and novice designers, *International*
1052 *Journal of Man-Machine Studies* 37 (1) (1992) 83–101.
- 1053 [41] G. Shanks, Conceptual data modelling: An empirical study of expert and novice
1054 data modellers, *Australasian Journal of Information Systems* 4 (2) (1997) 63–73.
- 1055 [42] J. Recker, Understanding process modelling grammar continuance: A study
1056 of the consequences of representational capabilities, Phd thesis, Queensland
1057 University of Technology (2008).
- 1058 [43] A. Maes, G. Poels, Evaluating quality of conceptual modelling scripts based on
1059 user perceptions, *Data & Knowledge Engineering* 63 (3) (2007) 769–792.
- 1060 [44] A. Bhattacharjee, Understanding information systems continuance: An
1061 expectation-confirmation model, *MIS Quarterly* 25 (3) (2001) 351–370.
- 1062 [45] J. Mendling, Detection and prediction of errors in epc business process models,
1063 Phd thesis, Vienna University of Economics and Business Administration
1064 (2007).

- 1065 [46] A. Gemino, Y. Wand, A framework for empirical evaluation of conceptual
1066 modeling techniques, *Requirements Engineering* 9 (4) (2004) 248–260.
- 1067 [47] D. Batra, J. A. Hoffer, R. P. Bostrom, Comparing representations with
1068 relational and eer models, *Communications of the ACM* 33 (2) (1990) 126–139.
- 1069 [48] K. Siau, P.-P. Loo, Identifying difficulties in learning uml, *Information Systems*
1070 *Management* 23 (3) (2006) 43–51.
- 1071 [49] J. Parsons, L. Cole, What do the pictures mean? guidelines for experimental
1072 evaluation of representation fidelity in diagrammatical conceptual modeling
1073 techniques, *Data & Knowledge Engineering* 55 (3) (2005) 327342.
- 1074 [50] J. Cohen, A coefficient of agreement for nominal scales, *Educational and*
1075 *Psychological Measurement* 20 (1) (1960) 37–46.
- 1076 [51] J. R. Landis, G. G. Koch, The measurement of observer agreement for
1077 categorical data, *Biometrics* 33 (2) (1977) 159–174.
- 1078 [52] S. Siegel, *Nonparametric Statistics for The Behavioral Sciences*, McGraw-Hill
1079 *Series in Psychology*, McGraw-Hill Kogakusha, Tokyo, Japan, 1956.
- 1080 [53] J. C. Nunnally, I. H. Bernstein, *Psychometric Theory*, 3rd Edition, McGraw-Hill
1081 *Series in Psychology*, McGraw-Hill, New York, New York, 1994.
- 1082 [54] C. Fornell, D. F. Larcker, Evaluating structural equations with unobservable
1083 variables and measurement error, *Journal of Marketing Research* 18 (1) (1981)
1084 39–50.
- 1085 [55] J. P. Stevens, *Applied Multivariate Statistics for the Social Sciences*, 4th
1086 Edition, *Applied Multivariate STATS*, Lawrence Erlbaum Associates, Hillsdale,
1087 New Jersey, 2001.
- 1088 [56] F. D. Davis, Perceived usefulness, perceived ease of use, and user acceptance of
1089 information technology, *MIS Quarterly* 13 (3) (1989) 319–340.

- 1090 [57] G. C. Moore, I. Benbasat, Development of an instrument to measure the
1091 perceptions of adopting an information technology innovation, *Information*
1092 *Systems Research* 2 (3) (1991) 192–222.
- 1093 [58] G. Alonso, M. Rosemann, P. Dadam (Eds.), *Business Process Management -*
1094 *BPM 2007*, Vol. 4714 of *Lecture Notes in Computer Science*, Springer, Brisbane,
1095 Australia, 2007.
- 1096 [59] M. Dumas, M. Reichert, M.-C. Shan (Eds.), *Business Process Management -*
1097 *BPM 2008*, Vol. 5240 of *Lecture Notes in Computer Science*, Springer, Milan,
1098 Italy, 2008.
- 1099 [60] E. Kindler, On the semantics of eps: Resolving the vicious circle, *Data &*
1100 *Knowledge Engineering* 56 (1) (2005) 23–40.
- 1101 [61] J. Mendling, *Metrics for Process Models: Empirical Foundations of Verification,*
1102 *Error Prediction and Guidelines for Correctness*, Vol. 6 of *Lecture Notes in*
1103 *Business Information Processing*, Springer, Berlin, Germany, 2008.
- 1104 [62] D. C. Sutton, Linguistic problems with requirements and knowledge elicitation,
1105 *Requirements Engineering* 5 (2) (2000) 114–124.
- 1106 [63] N. Guarino, Concepts, attributes, and arbitrary relations. some linguistic
1107 and ontological criteria for structuring knowledge bases, *Data & Knowledge*
1108 *Engineering* 8 (3) (1992) 249–261.
- 1109 [64] J. F. M. Burg, R. P. van de Riet, The impact of linguistics on conceptual
1110 models: Consistency and understandability, *Data & Knowledge Engineering*
1111 21 (2) (1997) 131–146.
- 1112 [65] P. J. Agerfalk, O. Eriksson, Action-oriented conceptual modelling, *European*
1113 *Journal of Information Systems* 13 (1) (2004) 80–92.
- 1114 [66] B. Kiepuszewski, A. H. M. ter Hofstede, W. M. P. van der Aalst, Fundamentals
1115 of control flow in workflows, *Acta Informatica* 39 (3) (2003) 143–209.

- 1116 [67] D. Gasevic, V. Devedzic, Petri net ontology, *Knowledge-Based Systems* 19 (4)
1117 (2006) 220–234.
- 1118 [68] J. Mendling, W. M. P. van der Aalst, Formalization and verification of eps with
1119 or-joins based on state and context, in: J. Krogstie, A. L. Opdahl, G. Sindre
1120 (Eds.), *Advanced Information Systems Engineering - CAiSE 2007*, Vol. 4495 of
1121 *Lecture Notes in Computer Science*, Springer, Trondheim, Norway, 2007, pp.
1122 439–453.
- 1123 [69] D. Hardt, Verb phrase ellipsis: Form, meaning, and processing, Phd thesis,
1124 University of Pennsylvania (1993).
- 1125 [70] G. Fliedl, C. Kop, H. C. Mayr, From textual scenarios to a conceptual schema,
1126 *Data & Knowledge Engineering* 55 (1) (2005) 20–37.
- 1127 [71] J. E. Ingvaldsen, J. A. Gulla, X. Su, H. Ronneberg, A text mining
1128 approach to integrating business. process models and governing documents, in:
1129 R. Meersman, Z. Tari (Eds.), *On the Move to Meaningful Internet Systems 2005:*
1130 *OTM Workshops*, Vol. 3762 of *Lecture Notes in Computer Science*, Springer,
1131 Agia Napa, Cyprus, 2005, pp. 473–484.
- 1132 [72] M. Ehrig, A. Koschmider, A. Oberweis, Measuring similarity between semantic
1133 business process models, in: J. F. Roddick, A. Hinze (Eds.), *Conceptual*
1134 *Modelling 2007*, Vol. 67 of *Conferences in Research and Practice in Information*
1135 *Technology*, Australian Computer Science Communications, Ballarat, Australia,
1136 2007, pp. 71–80.
- 1137 [73] T. Halpin, M. Curland, Automated verbalization for orm 2, in: R. Meersman,
1138 Z. Tari, P. Herrero (Eds.), *On the Move to Meaningful Internet Systems*
1139 *2006: OTM 2006 Workshops*, Vol. 4278 of *Lecture Notes in Computer Science*,
1140 Springer, Montpellier, France, 2006, pp. 1181–1190.
- 1141 [74] P. J. M. Frederiks, T. P. van der Weide, Information modeling: The process and

- 1142 the required competencies of its participants, *Data & Knowledge Engineering*
1143 58 (1) (2006) 4–20.
- 1144 [75] B. Nuseibeh, S. Easterbrook, Requirements engineering: A roadmap, in:
1145 C. Ghezzi, M. Jazayeri, A. L. Wolf (Eds.), 22nd International Conference on
1146 Software Engineering, ACM, Limerick, Ireland, 2000, pp. 35–46.
- 1147 [76] M. Rosemann, Preparation of process modeling, in: J. Becker, M. Kugeler,
1148 M. Rosemann (Eds.), *Process Management: A Guide for the Design of Business*
1149 *Processes*, Springer, Berlin, Germany, 2003, pp. 41–78.
- 1150 [77] C. Fillies, G. Wood-Albrecht, F. Weichardt, Pragmatic applications of the
1151 semantic web using semtalk, *Computer Networks* 42 (5) (2003) 599–615.
- 1152 [78] G. Guizzardi, H. Herre, G. Wagner, On the general ontological foundations of
1153 conceptual modeling, in: S. Spaccapietra, S. T. March, Y. Kambayashi (Eds.),
1154 *Conceptual Modeling - ER 2002*, Vol. 2503 of *Lecture Notes in Computer*
1155 *Science*, Springer, Tampere, Florida, 2002, pp. 65–78.
- 1156 [79] M. Hepp, F. Leymann, C. Bussler, J. Domingue, A. Wahler, D. Fensel, Semantic
1157 business process management: Using semantic web services for. business process
1158 management, in: *IEEE International Conference on e-Business Engineering*,
1159 IEEE, Beijing, China, 2005, pp. 535 – 540.
- 1160 [80] G. Wagner, The agent-object-relationship metamodel: Towards a unified view
1161 of state and behavior, *Information Systems* 28 (5) (2003) 475–504.
- 1162 [81] B. Levin, *English Verb Classes and Alternations: A Preliminary Investigation*,
1163 University of Chicago Press, Chicago, Illinois, 1993.
- 1164 [82] C. P. Menzel, M. Grninger, A formal foundation for process modeling, in:
1165 N. Guarino, B. Smith, C. Welty (Eds.), 2nd International Conference on Formal
1166 Ontology in Information Systems, ACM, Ogunquit, Maine, 2001, pp. 256–269.

- 1167 [83] J. Mendling, J. Recker, Towards systematic usage of labels and icons in
1168 business process models, in: T. Halpin, H. A. Proper, J. Krogstie (Eds.), 12th
1169 International Workshop on Exploring Modeling Methods in Systems Analysis
1170 and Design, CEUR Workshop Proceedings Series, CEUR, Montpellier, France,
1171 2008, pp. 1–13.
- 1172 [84] M. Boudreau, D. Gefen, D. Straub, Validation in information systems research:
1173 A state-of-the-art assessment, *MIS Quarterly* 25 (1) (2001) 1–16.
- 1174 [85] H. A. Reijers, J. Mendling, Modularity in process models: Review and effects,
1175 in: M. Dumas, M. Reichert, M.-C. Shan (Eds.), *Business Process Management*
1176 - BPM 2008, Vol. 5240 of *Lecture Notes in Computer Science*, Springer, Milan,
1177 Italy, 2008, pp. 20–35.
- 1178 [86] J. Mendling, H. A. Reijers, W. M. P. van der Aalst, Seven process modeling
1179 guidelines (7pmg), QUT ePrints Report 12340, Queensland University of
1180 Technology (2008).
- 1181 [87] C. B. Achour, C. Rolland, N. A. M. Maiden, C. Souveyet, Guiding use case
1182 authoring: Results of an empirical study, in: *IEEE International Symposium on*
1183 *Requirements Engineering*, IEEE, Limerick, Ireland, 1999, pp. 36–43.