HAL
open science

Multidisciplinarity vs. Multivocality, the case of "Learning Analytics"

Nicolas Balacheff, Kristine Lund

To cite this version:

Nicolas Balacheff, Kristine Lund. Multidisciplinarity vs. Multivocality, the case of "Learning Analytics". LAK 2013 - International Conference on Learning Analytics and Knowledge, Apr 2013, Leuven, Belgium. pp.5-13, 10.1145/2460296.2460299 . hal-00805362

HAL Id: hal-00805362

https://telearn.hal.science/hal-00805362v1

Submitted on 27 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ACM 2013 -- this is the author version of the work. It is posted here by permissin of ACM for your personal use. Not for redistribution. The definitive version will be published in "Proceedings of the third conference on Learning Analytics and Knowledge" (in press)

Multidisciplinarity vs. Multivocality, the case of "Learning Analytics"

Nicolas Balacheff
CNRS, LIG Research Lab
University of Grenoble
46 Avenue Félix Viallet 38031
Grenoble
+33 (0) 4 76 57 50 67
Nicolas.Balacheff@imag.fr

Kristine Lund
CNRS, ICAR Research lab
University of Lyon, ENS Lyon
15 parvis René Descartes 69700 Lyon
+33 (0) 4 37 37 63 16
Kristine.Lund@ens-lyon.fr

ABSTRACT
In this paper, we consider an analysis of the TeLearn archive, of the Grand Challenges from the STELLAR Network of Excellence, of two Alpine Rendez-Vous 2011 workshops and research conducted in the Productive Multivocality initiative in order to discuss the notions of multidisciplinarity, multivocality and interidisciplinarity. We use this discussion as a springboard for addressing the term "Learning Analytics" and its relation to "Educational Data Mining". Our goal is to launch a debate pertaining to what extent the different disciplines involved in the TEL community can be integrated on methodological and theoretical levels.

Categories and Subject Descriptors
H.1.0 [**Information Systems**]: Models and Principles – *General*.

General Terms
Measurement, Documentation, Human Factors, Standardization, Theory

Keywords
Technology Enhanced Learning, Learning Analytics, Educational Data Mining, Multidiscipinarity, Multivocality

1. THE TEL DICTIONARY PERSPECTIVE
The Technology Enhanced Learning (TEL) research area is by nature multidisciplinary, constantly importing terms, concepts and methods from the various disciplines participating in its development. TEL also coins terms, forges concepts and designs new methods in order to face the original problems it encounters. This evolution is so rapid and the motivations are so diverse that the language used is often not well defined. As a result, it is difficult to ensure that the wheel is not being reinvented despite what appears to be new research. In the course of TEL development, various terms may appear whose difference may be more related to variations in the communities that coined them rather than to fundamental differences in the concepts being considered. The term "Learning Analytics" is a good candidate for exploring these issues and for understanding the conceptual, linguistic and social stakes of the appearance of such a term and its associated challenges.

The term "Learning Analytics" was first used in 2009 by Bienkowski Mingyu & Means [3] and shortly thereafter, the first international conference on Learning Analytics was held in May 2011 in Banff (Alberta, Canada). Such a short delay suggests that the expression was not coined to respond to the need of a specific research program, but rather to identify the point of convergence of a community since holding a conference implies the existence of enough research and results to feed its program. This remark anticipates a question that we will raise about the nature of the term "Learning Analytics" — a term that could have one of two uses for a community. Either the community has identified a particular problem within its boundaries and the term aids in conceptualizing that problem, or the term was coined and has demonstrated its efficiency outside of the research community (e.g. at the frontier of another research community or to a general societal need) and is imported and plays a key role in defining a new community around its promises; this phenomena is often associated to a technology push (e.g. "mobile learning"). In the first case, if the community is single-discipline, then it is likely that researchers will use their tried and true methods to attempt to conceptualize and solve the problem. If the community is multidisciplinary, perhaps multiple frameworks and methods will be applied in parallel. If the term is imported from another research or practitioner community, there is an opportunity to integrate different frameworks and methods used towards a common goal, perhaps achieving a new vision or defining a new set of research questions.

We argue that the term "Learning Analytics" is currently mobilized within a multidisciplinary community of researchers and in this paper, we explore its origins and the opportunities this provides.

2. MULTIDISCIPLINARITY, INTERDISCIPLINARITY AND MULTIVOCALITY
Three concepts will play a key role in discussing the birth of the term "Learning Analytics" and its current status: multidisciplinarity, interdisciplinarity and multivocality. Van den Besselaar & Heimeriks [2] argue that if a disciplinary research field is defined as "a group of researchers working on a specific set of research questions, using the same set of methods and a shared approach (op. cit., p. 706)", then the different forms of so called non-disciplinary research such as multidisciplinarity, interdisciplinarity and transdisciplinarity are ways of combining elements from various disciplines in order to get these disciplines

(c) 2013 Association for Computing Machinery. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of the national government of France. As such, the government of France retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

LAK '13, April 08 - 12 2013, Leuven, Belgium

Copyright 2013 ACM 978-1-4503-1785-6/13/04…$15.00

to productively interact. Such interactions between disciplines can take many forms, ranging from communicating and comparing ideas, exchanging data, methods and procedures to mutually integrating concepts, theories, methodologies and epistemological principles (op. cit., 2001).

The claim here is that the indicator that distinguishes between the forms of non-disciplinary research is the level of integration of the disciplinary approaches that they are based upon. In van den Besselaar and Heimerik's view, neither theoretical perspectives nor actual results from different participating disciplines are integrated during multidisciplinarity. Rather, "the subject under study is approached from different angles, using different disciplinary perspectives (op. cit., p. 706)". Choi & Pak [5] hold a similar view, arguing that multidisciplinarity draws on knowledge from different disciplines, but each researcher group stays within its own boundaries. On the other hand, interdisciplinary research integrates contributing disciplines by creating its own theoretical, conceptual and methodological identity [2] or in other words, "analyzes, synthesizes and harmonizes links between disciplines into a coordinated and coherent whole ([5] p. 351)".

We agree with Leeds-Hurwitz [9], that disciplines are social constructions and that in order to carry out ones own research lucidly, it is necessary to be aware of disciplinary history, cognate disciplines, international variations and rival subdisciplines. The second author of this article and colleagues took up goals similar to those of Leeds-Hurwitz during a series of international workshops that involved sharing five corpora of group interactions and performing multiple analyses from different epistemological and methodological frameworks on each of them. We aimed for what we termed Productive Multivocality [10]. Multivocality makes reference to the different analytical "voices" that gathered around a particular corpus and those voices became "productive" when progress was made towards refining analytical concepts, rendering explicit epistemological positions, and in general characterizing under what conditions learning occurs in a set of corpora taken as being representative of the types of corpora studied in the Computer Supported Collaborative Learning (CSCL) community. These workshops were carried out during the following international conferences within a group of people interested in CSCL: International Conference on the Learning Sciences (ICLS 2008), Computer Supported Collaborative Learning (CSCL 2009), the STELLAR Alpine Rendez-Vous (ARV 2009), ICLS 2010 and ARV 2011.

We argue that multivocality is closer to interdisciplinarity than to multidisciplinarity. In order to illustrate this, let us take a closer look at the nature of the CSCL community in general. This closer look is intended to set the stage for reporting on the beginnings of a similar analysis of the communities laying stake to "Learning Analytics". Our underlying assumption is that multivocality and interdisciplinarity are approaches that move research fields forward. We argue that the communities researching "Learning Analytics" are nicely positioned to benefit from such approaches, much in the same way that CSCL has been.

In a study [17] that obtained answers from 15 out of 28 researchers who were representative of geographical locations and of levels of participation in CSCL from 1995 to 2005, the authors found that respondents worked from multiple disciplines: 15 people responded working from 21 disciplines (5 from Computer and/or Information Science, 4 from Education, 4 from Psychology (including 2 from Educational Psychology), 1 from Conversation Analysis and 1 from Knowledge Building. In another study that questioned reviewers from the 2007 CSCL conference, the disciplines mostly cited were Computer Science, Psychology, Educational Sciences and the Learning Sciences. However, a large variety of other disciplines were cited and thirty four disciplines were each cited by only one person. The granularity of these disciplines varied greatly going from large grain (i.e. Linguistics, Artificial Intelligence, Economics) to small grain (i.e. Curriculum and Methods of Teaching Arabic or Educational Measurement and Statistics). The field of CSCL is clearly multidisciplinary: different disciplinary perspectives are brought to bear upon the object of computer supported collaborative learning.

But is the CSCL community interdisciplinary or has a small part of the community made progress in that direction? As we have stated elsewhere [15] it is not possible nor desirable to attempt complete integration as the CSCL and larger Learning Sciences community are too diverse, both theoretically and methodologically. Indeed, diversity is one of the strengths that we wish to maintain since dialogues about analytical constructs between researchers that differ in their ontology and epistemology are particularly enlightening [1]. Instead of attempting to merge theoretical perspectives into a kind of 'super theory', we chose a set of *boundary objects* [14] to form the bridges between disciplines and create dialogue with the goal of better understanding how learning takes place in different contexts. An example of how the boundary objects we chose functioned as a bridge between disciplines will be presented in section 4.

Indeed, as the Productive Multivocality initiative culminates in publication of a book with Springer, expected in 2013, we claim to have achieved a certain level of integration between disciplines. We succeeded in that by sharing data with the supposition that this same data will serve the purpose of each different analyst's *problématique*[1]. If on the basis of this shared analytical object, we can claim an ontological convergence between the researchers who analyze the same corpus, we cannot claim a theoretical convergence, nor a methodological one. Multivocality is quite specific in that multiple voices are expected to identify key phenomena, but each voice will identify potentially different phenomena or if they are the same, they will specify them and explain them in different ways, while relying on different assumptions. It is only when these ways of explaining become explicit and are compared to other ways, with the shared analytical object serving as a boundary object, is it possible to approach either methodological convergence or theoretical convergence and thus interdisciplinarity.

This discussion of the extent to which multivocality may tend toward interdisciplinarity in the context of CSCL sets the stage for our analysis of the case of "Learning Analytics", although we have not yet the same kind of evidence.

In this communication we will first explore the lexical environment of "Learning Analytics", then we will discuss its scientific position in relation to the effort to define it, its disciplinary foundations and their potential interactions. Finally we will open the discussion concerning the scientific stakes of the emerging Learning Analytics community and its potential role in the TEL research area, at the crossroads of multivocality and multidisciplinarity.

---

[1] "Problématique" is a French word used to name a coherent set of problems and assumptions. It provides a coherent framework to express problems, why it is interesting to solve them and how the current research described is able to do so. This term is not a synonym for the English word "problematic".

3. "LEARNING ANALYTICS" AND RELATED TERMS AND CONCEPTS

From 2004 onwards, the TEL research community has made an effort to create a shared thematic open archive, TeLearn[2], where one can find a large number of articles and reports. In 2010, the first author did an analysis of the key terms and dominant expressions in this repository and it so happens that "Learning Analytics" was absent. This could mean that the community engaged in this area had not yet committed itself to contributing to the archive, but certainly we could at least conclude that its potential impact was not visible at that time. This is not surprising given its recent emergence. However, a search with the expression "Educational Data Mining" provided some results, as well as one with the expression "learning trails". This suggests a potential of interest for "Learning Analytics", indeed.

A more recent exploration of the outcomes of the 2011 Stellar Alpine Rendez-vous (ARV) and of the Grand Challenge problems as proposed by the Stellar network of excellence is more positive. It shows the appearance of "Learning Analytics" as a key word in the TEL research community and suggests relations to some other expressions. Figure 1 shows the clouds of keywords[3], organized according to their proximity in the texts of the Stellar Grand Challenge problems.

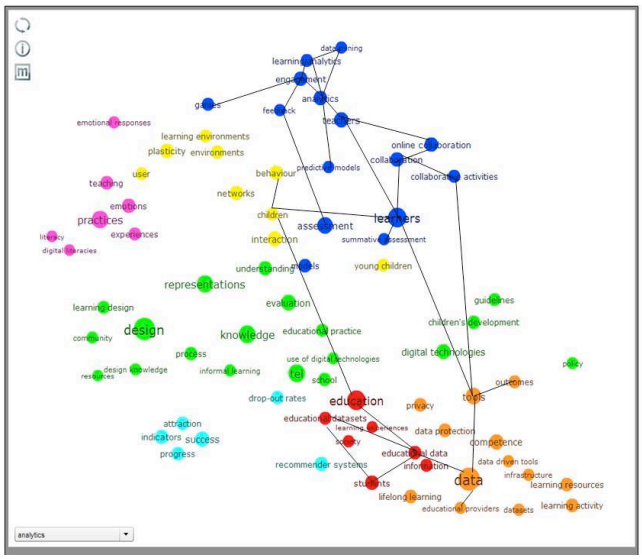

Figure 1. Keywords, organized by proximity in the texts of the Stellar Grand Challenge problems, and their relation to "Learning Analytics" (only partial links are shown to help readability). All links are available online[4]. Colors represent the clusters centered on the most important keywords.

[2] http://telearn.archives-ouvertes.fr/en

[3] This original map accounts for the co-occurrences present in this corpus. To build the map, we selected words that frequently appear in the same segment (the threshold is 50 words before - 50 words after the chosen word). If two words appear close to one another, they are linked in the map. Generic words (learning, technology) were deleted for this analysis as they brought too much noise. We present here a simplified version of the map so that it remains readable despite the limited space in this text.

[4] http://www.tel-thesaurus.net/maps/contexteGCP/GexfWalker/

Surprisingly, "data" and "Learning Analytics" have no direct connections in this particular representation, In order to unpack this lack of direct connections, we can look more closely at the structure of the graph; it shows that "Learning Analytics" is massively connected to "analytics" and "engagement" (cf. Figure 2). "Analytics" is connected to "games", "feedback", "predictive models", "teachers, "data mining" and "learners". The direct neighborhood of "data" is made up of 15 terms and expressions (cf. Figure 3): "tools", "competence", "privacy", "data driven tools", "data protection", "datasets", "educational data", "educational providers", "information", "infrastructure", "learning resources", "learning activity", "outcomes", "policy", "education".

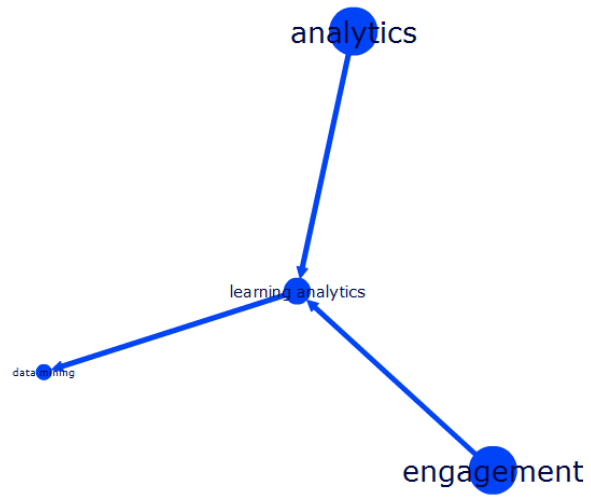

Figure 2. A close-up of the connections with "Learning Analytics" from Figure 1.

Figure 3. The terms closest to "data" in the Stellar Grand Challenges

As shown in Figure 1, the shortest path from "analytics" to "data" is via "learners" and "tools". While this makes sense without being very specific, it does not allow for differentiating between

"Learning Analytics" and "learning trails" and "Educational Data Mining", two expressions that are often associated to data in TEL Research. Therefore, we looked for another similar corpus that could help us differentiate between "Learning Analytics", "learning trails" and "Educational Data Mining". We decided to apply the same clouds of keywords methodology on all the white papers produced by the ARV 2011 workshops. Perhaps the small groups of close colleagues working on specific TEL issues could help us differentiate the expressions above.

When looking at the graph constructed on these white papers with the same methodology from the data extracted from the final texts of the Stellar Grand Challenge problems, one can observe that "analytics" has no connection to data or to dataset, arguably the two semantically closest keywords (cf. Figure 4.).

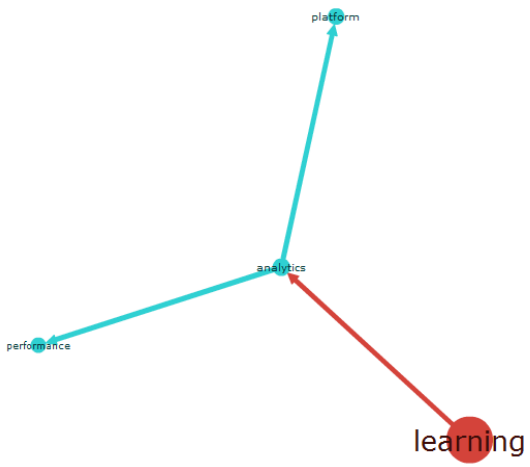

Figure 4. The terms closest to "analytics" in the white papers.

It appears very likely that this lack of connections between "data" and "Learning Analytics" or just to "analytics" would be due to a lack of connections between the research communities involved. The final work on the formulation on the Grand Challenge Problems was done in a larger context although it *began* with the white papers written by the ARV 2011 workshops so it could perhaps be understood that connections between "data" and "Learning Analytics" could be lost in that many researchers in the STELLAR network of excellence from different fields participated. But how is this also the case in the small groups of close researchers working together on specific TEL subjects? We therefore decided to look more closely at two workshops focusing heavily on "data": (ws8: Productive Multivocality and ws6: DataTEL). We will consider this issue more precisely in the next paragraphs.

We might wonder if the two workshops having "data" at the locus of their work are part of the same community. To what extent are they working with data in similar ways? Do they use similar vocabulary and have similar goals? If we analyze the white papers they provided after their workshops were over, we see that data is of different granularity and specific goals and vocabulary are different, although both workshops are interested in the high level goal of better understanding learning. Here, we compare the answers given by the two workshops to two questions involving defining Grand Challenge Problems for the future. The first question is "What problems of the European education system are addressed, and what are the long term benefits for society" and the second is "What are the main activities to address this Grand Challenge Problem?"

The first workshop "Productive Multivocality" answered the first question insisting on the variety of the approaches, on the nature of corpora and on empirical evidence: "*In order to make use of this variety of research data across the different research groups these relevant data sets should be shared and made accessible. The data sets and related analyses could serve as boundary objects and stimulate fruitful discussion across the different research approaches. This would not just show the multivocality in CSCL research, but could also serve as a means for converging evidence about the potentials and effectiveness of TEL and CSCL. This allows not just an overview about the effectiveness of CSCL in teaching and learning for researchers and the scientific community, but also for stakeholders and practitioners*".

The main activities needed to address this Grand Challenge Problem include the development of a technical infrastructure for supporting open data sharing and exchange of results and "lessons learned" among researchers, practitioners and stakeholders. The Productive Multivocality workshop also suggests the implementation and formative evaluation of a "*supportive structure for a dialogical interpretation of the data in order to make the community and stakeholders aware what results converge among the different data sets and different interpretations and in order to identify open questions*".

On the other hand, the DataTEL workshop answers the first question by claiming that "*The research on TEL recommender systems can contribute to decrease the drop-out rate [in education] by disseminating its research outcomes for the development of different support systems for teachers and students to offer relevant information at the right time"*.

The main activities needed to address this Grand Challenge Problem suggested by the DataTEL workshop are: *customize existing recommendation algorithms for learning*, employ recommender systems in real-life scenarios and develop suitable evaluation criteria for different kinds of recommender systems.

In comparing how these two workshops chose to define their Grand Challenge Problem, we note that although data is at the heart of both of their proposals, there is almost no shared vocabulary, apart from perhaps cognates of "learning". In addition, there is a great difference in terms of scope in the two workshop's objectives. The Productive Multivocality workshop aims to first produce an infrastructure for sharing data and second to elicit reasons for researchers from different disciplines to analyze this data. The DataTEL workshop centers on a specific problem that they aim to remedy — that of high drop-out rates in on-line education. Of course, it is hardly surprising that two different teams focus on different aspects of learning by analyzing data and that one focuses on solving a problem for practitioners (DataTEL) and the other on a problem designed to integrate research results from different communities (Productive Multivocality). It is clearly not enough for two research communities to hold an interest for data and to analyze data in order to better understand learning in order for them to share *problématiques*. However, as we will show in section 4 with the Productive Multivocality initiative, it is also not enough to share the *same* data and to analyze it to look for phenomena illustrating learning for researchers to share *problématiques*. Our intention in

this article is to use analyses of these other multidisciplinary contexts (i.e. TeLearn archive, STELLAR Grand Challenges, two Alpine Rendez-Vous 2011 workshops and Productive Multivocality) to show the potential for developing a mindset where theoretical and methodological approaches in Learning Analytics and Educational Data Mining are compared and contrasted with the goal of understanding each other's *problématiques*.

A first attempt at analyzing vocabulary used in the communities that are implicated in TEL was undertaken during the fall of 2011. This analysis is a precursor to understanding to what extent *problématiques* are already partially shared between participating disciplines and to pinpoint areas where perhaps existing tensions can be productive. A discussion was launched in the TEL Dictionary LinkedIn Group with the question of whether there was a specific concept underpinning "Learning Analytics" or whether it was a new flag in the TEL research community used as a shared sign of recognition among those researchers interested in exploring the benefit from importing an analytics *problématique* into TEL Research. This discussion didn't catch much attention, but left the idea that there was no clear view of how "Learning Analytics" could be differentiated from "Educational Data Mining".

In a recent effort to define "Learning Analytics", a seminal paper by Long and Siemens [10] refers to the definition proposed by the 1st LAK conference:

"*Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs*" (ibid. p. 34).

Let us compare it with the definition of "Educational data mining" prepared by Michel C. Desmarais and Ryan S.J.D. Baker for the TEL Dictionary:

"*Educational Data Mining is a term used for processes designed for the analysis of data from educational settings to better understand students and the settings which they learn in*." [5]

The distance between both definitions is not obvious for most TEL researchers. There is a general understanding that in order to design better adapted environments, one must be able to develop models and techniques to gather and to analyze data in order to make relevant pedagogical decisions and to provide feedback fulfilling learners' personal needs. Indeed, a recent US report expresses this shared concern: "*Two areas that are specific to the use of big data in education are* educational data mining *and* learning analytics*"* ([3] p. 8, authors' emphasis). The report points out an interesting difference which at first glance might appear as a difference of strategy: "*Unlike educational data mining, learning analytics does not generally address the development of new computational methods for data analysis but instead addresses the application of known methods and models to answer important questions that affect student learning and organizational learning systems*" [3].

We chose to use the classic method of exploring the origin and the related contexts of "Learning Analytics" and "Educational Data Mining" in order to better understand the possible difference between these two expressions. This is the core methodology of the construction of the TEL Dictionary; we question the history of the terms used by the involved research communities and trace the circumstances of their origin. From this perspective the two expressions have different profiles.

In the case of "Educational Data Mining", the birth of the expression and the respective community has a rather long history. The first workshops on data mining and learning [3] were held in the context of classic TEL conferences, namely AIED (2007, 2005), EC-TEL (2007), ICALT (2007), UM (2007), AAAI (2006, 2005), and ITS (2006, 2004, 2000). In 2008, the first conference specifically on Educational Data Mining was held in in Montréal. One will notice that Educational Data Mining has naturally taken its place within the flow of the history of data use in TEL research.

In the case of "Learning analytics", the origin is clearly exogenous. It has been inspired by the universe of business, characterized by data-driven decision-making and by business intelligence [3]. The objective is to take the tools and techniques which have already proven their efficiency in business and import them into TEL research so as to provide "*a new model for college and university leaders to improve teaching, learning, organizational efficiency, and decision making and, as a consequence, serve as a foundation for systemic change*." ([3], p. 32).

Learning analytics and Educational data mining are products of completely different processes and different *problématiques*. On this we might have a different analysis than Siemens and Baker [13]. These authors emphasize the focus of Educational Data Mining research on automatic processes, and that of Learning Analytics of providing information to stakeholders. Actually, designing open learner models is perfectly in line with the objectives of Educational Data Mining, and visualizing data after a relevant treatment is also an objective related to providing tools to support teachers or trainers. In both cases, learning as both a process and an epistemic outcome is targeted. On the other hand, Learning Analytics develops automatic modeling of massive data in order to provide relevant visualizations. But the key difference may not be located there. More important and significant, is the substantial difference of the epistemological grounds of both expressions so that potential conflicts are less justified by their conceptual *raison d'être* than by the difficulty in situating the respective communities within a larger research landscape. (Big) data is a shared field of action and confrontation of the two approaches could most likely be played out there. Within this aim, we submit the following questions to the communities involved:

- Are the Learning Analytics tools, techniques and strategies imported from the more general field of analytics sufficient for relevantly analyzing learning data?

- Should all data attached to the activities of a student be considered as learning data? (As opposed to strategies which are not related to learning but to the social management of the relations of the learner within his or her referent institution)

- Isn't Learning Analytics reducing successful learning to the academic success of students in their institutions, limiting de facto the *problématique* of TEL research?

- Compared to the classical *problématique* of "learner modeling" in AI, that of "learning trail analysis" and "Educational Data Mining", what are the specific contributions of "Learning Analytics"? Or is this comparison

[5] http://www.tel-thesaurus.net/wiki/index.php/Educational_data_mining

irrelevant? (in the case where the *problématique* is actually totally different).

Possibly for the first time in the TEL research area, the issue is not to address problems raised by multidisciplinarity, but those of multivocality and interdisciplinarity: that of the relations between different research communities sharing the same sources of evidence but not focusing on the same phenomenological observations, or if so, interpreting them in different frameworks.

In the following section, we use the Productive Multivocality context to illustrate how *problématique*s can converge in some ways between researchers who do not initially share an approach, either theoretically or methodologically. Our intent is to argue that this is an example of the kinds of opportunities available to "Learning Analytics" because it is — like CSCL — situated in a community made up of different disciplines.

4. HOW MULTIVOCALITY CAN TEND TOWARDS INTERDISCIPLINARITY

In this section we present a selection of results we obtained from sharing a corpus of Japanese 6th grade fractions within the Productive Multivocality initiative [15]. The results will show the difficulties involved in obtaining either methodological or theoretical convergence, but also specific instances of such convergence. Three researchers (Hajime Shirouzu, Ming Ming Chiu and Stefan Trausan-Matu) analyzed the corpus from their individual habitual perspectives. The first boundary object was thus the shared corpus itself, but we also defined a second boundary object. We asked the researchers to take "pivotal moments" into account during their analysis. At the start, such moments were not defined in a formal way, but there was a general and informal consensus that in the course of the learning process in a collaborative context, these moments correspond either to a rupture or to qualitative change. A pivotal moment is seen as a boundary object in the sense that its general and intuitive definition was understood, but it could be operationalized in different ways, depending on the researcher's *problématique*. So, whereas researchers agreed that learning was a temporal and dynamic process that corresponds to a change in state, depending on their *problématique*, they would characterize these states in different ways. They would also characterize the conditions for change in different ways. Note that a moment can also be an 'episode' because the temporal granularity changes with the perspective taken. In other words, *the unit of analysis is determined by the theoretical and methodological framework employed by the researcher*. Viewed in this way, the 'pivotal moment' becomes a lens with which one can analyze the divergence or the convergence between the analysts' approaches. We asked the researchers to look for 'pivotal moments' within a shared corpus provided by one of them. In the next three sections, we first give an extract from the corpus and an example of a particular moment that was designated as pivotal by each researcher. Second, we show how different voices can interpret the shared data without any type of convergence. Finally, we give an example where methodological and/or theoretical convergence occurs, thus tending towards interdisciplinarity.

4.1 An example of a pivotal moment designated by each researcher

In the figure below, we show a stretch of talk in which the three researchers who analyzed the corpus of fractions in a Japanese 6th grade classroom, each determine that a particular moment (of slightly different temporal lengths) is pivotal, but for different reasons that are linked to their *problématiques*.

The corpus that was recorded involved six Japanese 6th grade children studying fractions in one classroom session where the question the teacher asked was " Can you cut ¾ of 2/3 of a piece of origami paper?" [12]. After the children had all folded the origami paper in different ways, in order to attempt the answer the question, they were asked the question "Are the answers the same?". It is during this questioning that the extract in Table 1 occurs.

Table 1. A stretch of talk in which the three researchers each found a pivotal moment; each text type (i.e. bold, italic, underlined) corresponds to a researcher's definition of pivotal moment.

470	Y	**[Moves toward the teacher's desk by further raising his hip]**
471	Anon	[Whispers] The shapes **differ**
472	Y	*Differ. [with clear voice]*
473	Y	*Though areas are the same [with low voice]*
474	G	*The areas are* **the same**
475	T	yes
476	G	<u>but the shapes and production methods</u> **differ.**

In this short extract, we first see that student Y moves toward the teacher's desk by raising his hip. The video at this point in time shows that student Y is the last student out of six to enter into the learning space (i.e. to pay attention to the origami sheets being folded and discussed and to what is being written on the chalkboard). Looking for differential inter-animation patterns within a Bakhtinian perspective [16] (his pivotal moment is in bold in Figure 3), Trausan-Matu analyzes Y's behavior as being *characteristic of a divergent thinker and it is for him a pivotal moment in and of itself.* In addition, apart from the anonymous student at line 471, it is student Y who clearly announces that the shapes differ and this is the beginning of the pivotal moment that continues to line 476. At line 474, student G says the areas of the shapes are the same, but that the shapes and production methods differ (line 476). This comparison of characteristics of the origami paper, some of which are the same while others are different present an opportunity for the students to learn (i.e. change their thinking about the relationship between area and shape).

Shirouzu was looking in part for moments that helped him develop his own theory of focus-based constructive interaction, based on Miyake [11]. His questions concerned why students chose to focus on a certain aspect of their folding activity (e.g. shapes or production methods) during the interaction and where the interaction went as a result. In the extract in Figure 3, Shirouzu defines a pivotal moment at lines 472-474 (in italics), where he considers that one student focuses on the shapes differing whereas the other focuses on the fact that the areas are the same. Shirouzu recognizes that the students are speaking about the different characteristics of the origami paper, but although he is asking if there is collective understanding about the relationship between area and shape), he is more focused on *how the individual foci came about and how they were now going to influence where the interaction would go*.

Chiu's pivotal moments were breakpoints that illustrated a change in the quality of the interaction. He used his own method

called Statistical Discourse Analysis and analyzed how characteristics of recent turns of talk such as questions and evaluations are linked to characteristics of subsequent turns of talk, such as correct ideas, new ideas or justifications. For this extract, he identified a pivotal moment at line 476 (underlined), after which there was a sharp increase in new ideas.

Shirouzu's and Trausan-Matu's pivotal moments are sequences of conversation turns whereas Chiu's method restricts his pivotal moments to one conversation turn (requiring further qualitative analysis to identify a pivotal moment's boundaries). Whereas Shirouzu's and Trausan-Matu's pivotal moments focus on conceptual thinking, Chiu's pivotal moments cover the entire classroom interaction. Chiu's premise is that when students are asked to solve a new problem, they try to create new ideas (termed micro-creativity) and they assess their utility via explanation or justifications [4]. One principal question is how classroom processes affect new ideas and justifications and whether their effects differ across time. Pinpointing students' justifications allows the researcher to locate where students may be arguing about the relationship between area and shape and thus potentially changing their thinking.

4.2 Multivocality without convergence

In this section we show how the comparison of Shirouzu's, and Chiu's pivotal moments lead to progress in each other's *problématiques*, but not to integrating on either a theoretical or methodological level. However, the following discussion illustrates how one researcher's methods can be mobilized, once another researcher's goals are understood, even if the method is not subsequently appropriated by the second researcher.

Chiu performed new analyses focused on the class discussion activity phase of the pedagogical task after understanding that Shirouzu had a special interest in it. Shirouzu demonstrated that he was able to match new meanings to Chiu's interpretations of pivotal moments (occurring in Chiu's framework) that were relevant to him in his own framework. For example, Shirouzu saw his first pivotal moment as a collective display of new understanding whereas Chiu viewed it as indicating the end of a period of frequent ideas, occurring just after teacher acknowledgment. Indeed it is compatible that the moment when collective understanding is reached could correspond to the beginning of a drop in new ideas because learners are consolidating their knowledge in terms of concepts already expressed. Re-examining this moment in terms of Chiu's definition of ideas as "new" or "old" led Shirouzu to suggest that in his own framework, new ideas could correspond to conceptual or procedural changes of how to view the solutions, progressing potentially towards a collaborative pivotal moment. Shirouzu also noticed that Chiu's five breakpoints corresponding to frequency of new ideas also corresponded to when and how the pedagogical designer's intentions were actualized by students' behavior.

The previous discussion shows that when researchers compare the moments that they consider pivotal for learning, and one of them (Shirouzu) discovers that a pivotal moment he did not consider initially as pivotal is considered as being pivotal in another researcher's framework (Chiu's), he is capable of finding a reason why those moments could also be meaningful in his own framework. *In this case, neither methodological nor theoretical convergence is achieved, but a discussion has begun*.

4.3 Multivocality with convergence

In this section we show how the comparison of Shirouzu's and Trausan-Matu's pivotal moments lead to progress in Trausan-Matu's own *problématique*, but also to integrating both reearchers' approaches on a methodological level.

Trausan-Matu has used a semi-automatic content-based analysis system PolyCAFe (Polyphonic Conversation Analysis and Feedback generation) for the analysis of chat logs taken from collaborative learning sessions (Trausan-Matu and Rebedea, 2010). The Productive Multivocality collaboration introduced him to the analysis of transcribed oral conversations with both talk and gesture, a type of corpus he had not focused on before. Adding gesture to his analysis of human interaction amounts to extending the domain of application of his tool but more importantly to extending the concept of Bakhtin's 'voices' to include gestures. We interpret this re-conceptualization of 'voices' to mean that when Trausan-Matu was confronted with a corpus that presented forms of interaction he was not used to analyzing (i.e. gestures), he was able to *re-consider the types of data he took into account as important for understanding learning and to integrate them into his theoretical and methodological framework*. This change in conceptualization illustrates how closely related our theoretical frameworks are to the nature of the data we analyze. We argue that convergence occurs here in that Trausan-Matu widened Bakhtin's framework in order to take into account new types of corpora and by doing so, came closer to Shirouzu's framework.

Our final example illustrates the possibility of further methodological and theoretical integration (although this has not yet occurred). We could imagine Shirouzu using Trausan-Matu's tool in order to locate moments of personal foci (e.g. shape, area or production methods) since if the tool can locate inter-animation patterns and if differential patterns of concepts (i.e. shape versus area) can be considered as being opposed in some way, they are also differences in focus. There is thus possibility for methodological integration between Shirouzu's and Trausan-Matu's *problématiques*. In addition, on a theoretical level, conceptual differential patterns and personal foci (i.e. those having to do with, for example shape and area and not just sequences of "yes, no, yes, no") can be considered as converging analytical concepts.

In the next and final section, we take this discussion of multivocality that tends towards divergence or that tends towards convergence (and therefore towards interdisciplinarity in the latter case) and we transfer it to another context, that of using "Learning Analytics" as a boundary object whose study may enhance collaboration between technology and education.

5. "DATA" AS A BOUNDARY OBJECT FOR LEARNING ANALYTICS AND EDUCATIONAL DATA MINING

As an expression, "Learning analytics" can be understood in a radical way as the concatenation of "learning" and "analytics", the former indicating an objective and the latter a means: using analytics to improve learning. It is an interesting case since in most cases the technology push comes from the hard(ware) side but this time, it comes from the soft(ware) side. The reason why this borrowing of methods and approaches from outside of TEL research is considered promising is because their efficiency has been acknowledged in other areas that potentially share with TEL research only the emergence of Big Data about users. But the type of data is not clear and the use of such a generic word induces a spurious consensus. Indeed, what is there in common between the

learning trail, tracks, or traces of a learner or a group of learners struggling with learning the multiplication of ratios, and the log file of students using a LMS for a long period in different disciplines and for different purposes? Or, what is shared by a teacher having to manage a lesson on ratio and proportion, and by the dean of the university having to ensure the success of the freshman class? Indeed, there is an objective shared goal at stake: the success of the learners, but the data needed to answer these questions is very different. In addition, the problems raised are of a rather different nature and the meaning of "learning" in both cases is significantly different. This raises questions whose responses do not rest in the technology of analytics, but in the capacity of researchers from different origins and with different *problématiques* to cooperate constructively, like the researchers in our Productive Multivocality examples. Siemens and Baker [13] invite researchers from Educational data mining and Learning analytics to cooperate following an analysis of both research approaches. Such an invitation deserves a serious discussion for which we do not have the space in this short communication. For the purposes of this article, Educational data mining is claimed to be limited to the search for automated methods and Learning Analytics is claimed to be characterized by its objective to provide resources for human judgment. This is surely a much too rapid analysis or may likely even be misleading. Learning Analytics should become an object to be discussed and questioned from the perspective of the different *problématiques* interested in taking up the challenge of its importation into the Learning Sciences. In other words, it is time for TEL researchers to go beyond the soft(ware) push and appropriate a meaning for the tools of analytics from an educational and learning perspective. The agenda for making this move will include working on the same data from different perspectives and confronting the different understandings of Learning Analytics, without necessarily choosing any one particular definition as the one that is destined to become canonical. Finally, this type of exercise will allow researchers to consider to what extent the different disciplines involved can be integrated on both methodological and theoretical levels.

In order to reach this goal, we should first be more specific about what we mean by data in the contexts where Learning Analytics and Educational Data Mining are used. Why is data a potential boundary object for both domains? The Productive Multivocality context illustrated that being exposed to new types of data led to the extension of a theoretical and methodological framework and thus to a widened capability of analysis. If Learning Analytics focuses on simple timestamped online forum data, or on replayable and synchronized traces of multi-source collaborative activity [6] or on characteristics of students gleaned from questionnaires or institutional data, or finally on the informal ways of communicating (e.g. Facebook, Twitter, blogs) or around formal content (e.g. Moodle course), it seems reasonable to assert that the theoretical and methodological frameworks used to perform Learning Analytics or to do Educational Data Mining in those cases would be different. In order to understand the reasons for that, we need also to look at the *problématique* of the researchers involved to evaluate the potential for convergence in the community. What are their assumptions about data? About learning? What specific problems are they trying to solve? Why is it interesting to solve them? For example, how can different approaches such as Social Network Analysis, Discourse Analysis and Multimodal Analysis — each focusing on a particular aspect of collected data, and each situated within a particular perspective — inform each other on a theoretical level?

The Productive Multivocality context also illustrated that a first researcher's methods used on a second researcher's data can lead to new insights for the second researcher even if (s)he does not appropriate the research method itself. This is another example of the strength of sharing data and succeeding in understanding each other's *problématiques* and is more straightforward and more common than the previous example.

Researchers who study group interactions — like those in the Productive Multivocality context, are similar to researchers who do Learning Analytics and Educational Data Mining in that they work in multidisciplinary communities where *problématiques* are different. We join Siemens and Baker [13] in a call for cooperation between these two domains with the suggestion of an analysis of the nature of data and of *problématiques* in order to explore the possibilities and potential advantages of partial convergence, perhaps in the name of a new theoretical, conceptual and methodological identity for both Learning Analytics and Educational Data Mining.

6. ACKNOWLEDGMENTS

We thank the TEL community for contributing to the TEL dictionary and thesaurus as well as the leaders of the ARV 2011 workshops Productive Multivocality (Carolyn Rosé, Kris Lund, Dan Suthers, Gregory Dyke) and DataTEL (Hendrik Drachsler, Katrien Verbert, Miguel-Angel, Martin Wolpers, Nikos Manouselis, Riina Vuorikari and Stefanie Lindstaedt. In addition we thank the analysts of the Japanese fractions corpus (Hajime Shirouzu, Stefan Trausan-Matu and Ming Ming Chiu). Our thanks also go to Emilie Manon and Jérôme Zeiliger who carried out the treatment needed for the corpora of texts we use here.

7. REFERENCES

[1] Abend, G. (2008). The Meaning of 'Theory'. Sociological Theory. Vol. 26, No. 2, pp. 173-199.

[2] van den Besselaar, P. and Heimeriks, G. 2001., Disciplinary, Multidisciplinary, Interdisciplinary: Concepts and Indicators. In M. Davis and C.S. Wilson (Eds), *ISSI 2001, 8th international conference of the Society for Scientometrics and Informetrics*, Sydney: UNSW 2001. pp. 705-716.

[3] Bienkowski, M., Feng, M. and Means, B. (2012). Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics, Department of Education's, Office of Educational Technology, http://ctl2.sri.com/eframe/wp-content/uploads/2012/04/EDM-LA-Brief-Draft_4_10_12c.pdf PDF (Draft version April 12 2012).

[4] Chiu, M. M. (2008). Effects of argumentation on group micro-creativity. *Contemporary Educational Psychology*, 33, 382-402.

[5] Choi, B. C. K. & Pak, A. W. P. 2006. Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. *Clinical and Investigative Medicine. Medecine Clinique et Experimentale*, 29(6), 351-364.

[6] Dyke, G., Lund, K., & Girardot, J.-J. (2010). Tatiana : un environnement d'aide à l'analyse de traces d'interactions humaines. *Technique et Science Informatiques, 29*(10), pp.1179-1205.

[7] Jaffe, E. 2009. Crossing Boundaries: The Growing Enterprise of Interdisciplinary Research. *Observer* Vol.22, No.5 May/June, 2009**.**

[8] Kienle, A. & Wessner, M. 2006. The CSCL Community in its First Decade: Development, Continuity, Connectivity. *International Journal of Computer-Supported Collaborative Learning (ijCSCL)* 1 (1), pp. 9- 33.

[9] Leeds-Hurwitz, W. 2012. These fictions we call disciplines. *The Electronic Journal of Communication / La Revue Electronic de Communication* Volume 22 Numbers 3 & 4, 2012

[10] Long, P. & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review Online*, 46, 5, pp. 31-40.

[11] Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive Science*, 10, 151-177.

[12] Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. Cognitive Science, 26, 469-501.

[13] Siemens G., Ryan S J.d. Baker 2012. Learning Analytics and Educational Data Mining: Towards Communication and Collaboration, *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, Pages 252-254, ACM: New York, NY, USA.

[14] Star, S. L., & Griesemer, J. R. 1989. Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology. Social Studies of Science, 19(3), 387-420.

[15] Suthers, D. D., Lund, K., Rose, C., Dyke, G., Law, N., Teplovs, C., Chen, W., Chiu, M., Jeong, H., Looi, C-K., Medina, R., Oshima, J., Sawyer, K., Shirouzu, H., Strijbos, J-W., Trausan-Matu, S. & van Aalst, J. (2011). Towards productive multivocality in the analysis of collaborative learning. In H. Spada, G. Stahl, N. Miyake, N. Law & K. M. Cheng (Eds.), Connecting Computer-Supported Collaborative Learning to Policy and Practice: Proceedings of the 9th International Conference on Computer-Supported Collaborative Learning (CSCL 2011), (Vol. III, pp. 1015-1022). Hong Kong: International Society of the Learning Sciences.

[16] Trausan-Matu, S., & Rebedea, T. (2009). Polyphonic Inter-Animation of Voices in VMT, in Stahl.G. (Ed.), Studying Virtual Math Teams (pp. 451 - 473). Boston, MA: Springer US.

[17] Kienle, A., Wessner, M., (2006). The CSCL Community in its First Decade: Development, Continuity, Connectivity. *International Journal of Computer-Supported Collaborative Learning (ijCSCL),* Vol. 1, No. 1, pp. 9-33.