On the Effort Required by Function Point Measurement Phases

Luigi Lavazza
Dipartimento di Scienze Teoriche e Applicate
Università degli Studi dell'Insubria
Varese, Italy
email: luigi.lavazza@uninsubria.it

***Abstract*— Function Point Analysis is widely used, especially to quantify the size of applications in the early stages of development, when effort estimates are needed. However, the measurement process is often too long or too expensive, or it requires more knowledge than available when development effort estimates are due. To overcome these problems, early size estimation methods have been proposed, to get approximate estimates of Function Point measures. In general, early estimation methods adopt measurement processes that are simplified with respect to the standard process, in that one or more phases are skipped. Early estimation methods are considered effective; however there is little evidence of the actual savings that they can guarantee. To this end, it is necessary to know the relative cost of each phase of the standard Function Point measurement process. This paper presents the results of two surveys concerning the relative amount of effort required by the phases of the standard Function Point measurement process. The analysis of the collected data can be used to assess the expected savings that early estimation methods make possible.**

Keywords- functional size measurement; Function Point Analysis; IFPUG Function Points; Simple Function Point; measurement process; cost of measurement; measurement effort.

I. INTRODUCTION

Knowing the relative cost of the phases that compose Function Point Analysis (FPA) is of great importance for software project managers [1].

FPA [2][3][4][5] is widely used. Among the reasons for the success of FPA is that it can provide measures of size in the early stages of software development, when they are most needed for cost estimation.

However, FPA performed by a certified Function Point (FP) consultant proceeds at a relatively slow pace: between 400 and 600 FP per day, according to Capers Jones [6], between 200 and 300 FP per day according to experts from Total Metrics [7]. Consequently, measuring the size of a moderately large application can take too long, if cost estimation is needed urgently. Also, the cost of measurement can be often considered excessive by software developers. In addition, cost estimates may be needed when requirements have not yet been specified in detail and completely.

To overcome these problems, early estimation methods (EEM's) have been proposed: these methods provide approximate estimated values of FP measures. Instead of going through the standard FP measurement process, EEM's provide estimates based on a little number of parameters that can be collect in a short time and with little effort. Most EEM's estimates are obtained via statistical models and contain unavoidable estimation errors. A list of EEM's can be found in [8] and [20].

The goal of the work presented here is to assess the cost of the measurement activities in terms of the effort required. However, as mentioned in the introduction, there is little agreement on the effort needed to carry out FP measurement: for instance, Capers Jones [6] and Total Metrics [7] provide quite different evaluations. Therefore, it appears more feasible to pursue an evaluation of the *relative* effort required by the measurement phases (i.e., the fraction of the total measurement effort dedicated to each phase). In this way, we can assess how much we save –in terms of measurement effort, hence ultimately of money– by skipping a measurement phase, i.e., by not performing one of the activities of the standard measurement process. In fact, if a manager knows that applying the standard measurement process in her organization takes X PersonHours per FP, and a simplified measurement process allows for saving 30% of the effort, she can easily conclude that in her organization the application of the simplified process will take 0.7X PersonHours. Of course, our manager should also take into account that Early Estimation Methods provide *estimates* of the actual measures, that is, the savings are usually associated to a loss of accuracy (not dealt with in this paper: interested readers can find some evaluations in [20]).

To estimate the relative effort required to carry out each phase of the FP measurement process, we made two surveys. Both surveys involved collecting opinions from expert measurers concerning the relative effort required by measurement phases. The first survey was carried out completely on-line, using a web-based questionnaire management system. The second one was carried out at a conference: the questions were illustrated to the audience and answers were provided on paper at the end of the conference.

The paper is structured as follows. Section II reports a few basic concepts of FPA. Section III describes how the first survey was carried out and illustrates the results of the survey. Section IV is similar to Section III, but it is dedicated to the second survey. In Section V we discuss how the results given in Sections III and IV can be used to assess the savings that can be obtained by using EMM's. Section VI discusses the threats to the validity of this study. Section VII accounts for related work. Finally, Section VIII draws conclusions and briefly sketches future work.

II. FUNCTION POINT ANALYSIS CONCEPTS

FPA aims at providing a measure of the size of the functional specifications of a given software application.

A. *The model of the software being measured according to FPA*

FPA addresses functional specifications that are represented according to a specific model. The model of functional specifications used by FPA is given in Figure 1. Briefly, Logical files are the data processed by the application, and transactions are the operations available to users. The size measure in FP is computed as a weighted sum of the number of Logical files and Transactions. The weight of logical data files is computed based on the Record Elements Types (RET), i.e., subgroups of data belonging to a data file, and Data Element Types (DET), i.e., the elementary pieces of data; besides, the weight depends on whether the data file is within the boundaries of the application, i.e., it is an Internal Logic File (ILF) or it is outside such boundaries, i.e., it is an External Interface File (EIF). The weight of transactions is computed based on the Logical files involved –see the FTR (File Type Referenced) association in Figure 1– and the DET used for I/O; besides, the weight depends on the "main intent" of the transaction. In fact, depending on the main intent, transactions are classified as External Inputs (EI), External Outputs (EO) or External Queries (EQ).

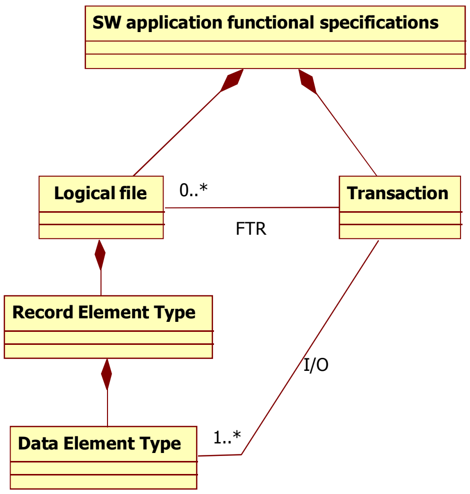

Figure 1. The model of software used in FPA.

B. *The FPA measurement process*

According to the International Function Point User Group (IFPUG) measurement manual [4][5], the measurement process includes the following phases:

1. Gathering the available documentation concerning functional user requirements;
2. Identifying application boundaries and determining the measurement goal and scope;
3. Identifying Elementary Processes (Transactions) and Logical Data Files;
4. Classifying transactions as EI, EO or EQ; classifying files as ILF or EIF; identifying RET, DET, FTR and determining complexity;
5. Calculating the functional size;
6. Documenting and presenting the measurement.

C. *Early Estimation methods*

EEM's tend to skip as many as possible of the steps listed above. The idea is straightforward: the less phases have to be performed, the faster and cheaper is the process. However, some activities –e. g., those involved in phases 1 and 2– are preparatory of the real measurement and cannot be skipped. Similarly, phase 6 can hardly be avoided. In any case, it should be noted that the simplification of the measurement process can affect phases 1 and 6 as well: on the one hand, a simplified process requires less documentation concerning the functional specifications of the application; on the other hand, documenting and presenting a simplified measurement is easier and faster than documenting the full-fledged measurement.

As a final observation, the extent of phase 6 depends on the context and the goal of measurement: for instance, if an organization is measuring the size of the application to be developed for internal purposes, the documentation can be kept to a minimum; on the contrary, if the functional size measures have to be used in a bid or in establishing the price of a contract, the documentation to be produced has usually to be quite detailed, and the presentation of the measures and measurement has also to be accurate.

In conclusion, EEM's address mainly phases 4 and 5. However, there is hardly any evidence of how much you save if you skip or simplify any of these phases. On the contrary, some evidence exists that by simplifying the measurement process, some measurement error is introduced [20].

D. *Other functional size measurement methods*

Besides IFPUG function points, several other methods have been proposed. These are treated as follows, in this paper:

- NESMA (Netherlands Software Metrics Association) function points [10] have become essentially equivalent to IFPUG Function Points. Therefore, the analysis presented in the paper applies to NESMA FP measurement as well.
- The Simple Function Point (SiFP) method [16][17][26][27] adopts a model of the software to be measured that is greatly simplified with respect to the model given in Figure 1. Hence, the measurement is similarly simplified: it requires only that transactions and logical data files are identified; then the functional size is computed as follows:

  $$\text{Size} = 4.7\ \#UGEP + 7\ \#UGDG \qquad (1)$$

  where #UGEP is the number of transactions (UGEP is for Unspecified Generic Elementary Process) and #UGDG is the number of logical data files (UGDG is for Unspecified Generic Data Group). Although it adopts a simplified measurement process, the SiFP method is not an EEM method: it is a proper functional size measurement method, which provides a proper measure (expressed in SiFP) [18], not an approximate estimate of the size expressed in IFPUG FP. The outcomes of our analysis can be used to evaluate the effort required for the SiFP process in comparison with the IFPUG measurement process.

- Other methods (like the COSMIC method [29], MKII Function Points [30] of FISMA [28]) are not considered in this paper, and the results we present are not applicable to these methods.

As a final remark, we remind that Function Point can be adjusted or not adjusted. In this paper we consider only unadjusted Function Points. This is consistent with the choice of the International Standardization Organization, which standardized unadjusted FP, not adjusted FP.

III. THE FIRST SURVEY

As mentioned in the introduction, the study reported here is based on two surveys. In this section we report about the first survey.

A. *The survey*

The investigation described here was performed via a questionnaire, which was filled by people that are experienced in IFPUG Function Point measurement.

The questionnaire was published on the kwiksurveys site [21]. The questionnaire was publicized via several channels:

- An invitation to fill out the questionnaire was sent to the Italian Software Measurement Association (www.gufpi-isma.org);
- A similar invitation was sent to the Nesma association [22];
- Finally, a question was published on ResearchGate [23], and experts were redirected to the questionnaire URL.

The questionnaire is reported in Appendix A. It can be noticed that the questionnaire targets both the IFPUG [4][5] and the Nesma [10] measurement processes. In fact, according to Nesma, "*[Since 1994,] owing to [...] the intensive cooperation between the Nesma and the IFPUG, the counting guidelines of the NESMA and the IFPUG continuously came closer and closer. [...] With the publication of IFPUG CPM 4.2 (2004) the last major differences between IFPUG and NESMA disappeared.*" Therefore, mixing data concerning the current IFPUG and Nesma measurement processes is perfectly safe, and the results found apply equally well to both measurement methods.

The questionnaire was published in November 2014, and answers were collected until April 2015.

B. *The Results of the survey*

31 answers were collected. Even if the number is not very large, it is nonetheless sufficient to get a reasonably reliable assessment of the relative effort required by FP measurement activities.

Of the respondents, 21 are certified Function Point Specialist (CFPS), and 4 are certified Function Point Practitioner (CFPP). Only 6 have no certification; however, of these, 2 use NESMA Function Points, therefore it is reasonable that they do not need an IFPUG certification.

The experience of the respondents is also quite reassuring: 20 respondents have been using FP measurement for over 10 years; only two for less than 5 years.

It should be noted that the questionnaire does not ask for a specific percentage for each phase; instead, it asks to specify in what range the actual percentage of effort belongs. This choice was due to two reasons: 1) the free version of the questionnaire provided by kwiksurveys does not support the collection of numeric values, and 2) it is unlikely that a respondent knows the exact fraction of effort that is spent in each phase, while it is much more probable that he/she can indicate the correct range.

The collected data concerning the relative effort required by each measurement phase are given in Table VI in Appendix A.

When information is collected via questionnaires, it is always possible that some respondents do not provide correct data. Therefore, before proceeding to the analysis of the collected data, it is necessary to remove unreliable answers from the dataset. In our case, the following problems were detected:

1) The sum of the efforts spent in each phase must be 100%. Having asked for ranges, we expect that the sum of the lower bounds of the ranges is ≤ 100% (but close to 100%) and that the sum of the upper bounds is ≥ 100% (but close to 100%). Respondents 12, 23 and 31 do not satisfy these conditions: total effort is in [27%, 60%] range for respondent 12, in [200%, 230%] range for respondent 23 and in [12%, 45%] range for respondent 31. These are clearly worthless indications, therefore they have been excluded from the dataset.
2) Among the remaining respondents, it is easy to spot a few outliers. Respondent 19 declared a fraction of effort for phase 6 (Documenting and presenting the measurement) that is almost half the total effort and more than double than the other respondents'. Respondent 27 declared an abnormally large amount of effort dedicated to phase 1 (Gathering the available documentation concerning functional requirements): such a large effort may be required in specific contexts, but is not representative of the general case (as other respondents clearly show). To preserve the representativeness of the data, the answers provided by the mentioned respondents have been excluded from the dataset.
3) Respondents 4 and 5 declared that they use EEM's. Their answers were removed from the dataset, since we are interested in the relative effort for the phases of the *standard* measurement process.

The answers provided by respondents concerning each measurement phases are summarized in Figure 2, where the boxplots of relative effort –expressed as a percentage– are given (diamond-shaped points indicate the mean values).

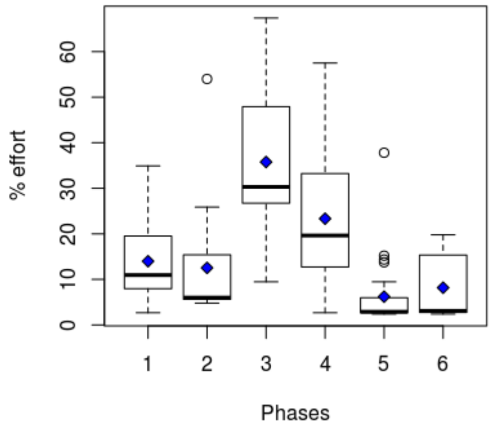

Figure 2. Distribution of relative phase efforts for the first survey dataset.

To analyze the data in Table VI, the following procedure was adopted:

1) Let ML_{ij} and MU_{ij} be the minimum and maximum relative effort for phase i indicated by respondent j. First of all, we computed for every phase i and for every respondent j, $M_{ij} = (ML_{ij} + MU_{ij})/2$: M_{ij} is the most likely effort for phase i according to respondent j.
2) For every respondent j, we computed $T_j = \Sigma_{i=1,6}\ M_{ij}$. Of course, we would like that $T_j = 100\%$, that is, the sum of the relative effort spent for each phase should be the total effort spent for measurement. In general it was not so. Thus, we normalized each phase effort: we computed $WM_{ij} = 100\ M_{ij}/T_j$. So, $\Sigma_{i=1,6}\ WM_{ij} = 100\%$.
3) For every phase i, we computed $M_i = (\Sigma_{j\in J}\ WM_{ij})/|J|$, where J is the set of respondents. M_i is the average effort indicated by respondents for phase i.

The resulting values (expressed as percentages) are given in Table I.

TABLE I. MEAN AND MEDIAN VALUES OF (NORMALIZED) PHASE RELATIVE EFFORTS

Phase	1	2	3	4	5	6
Mean	14.0	12.5	35.8	23.3	6.2	8.2

The results of the analyses provide some useful indications concerning the relative effort required by the phases of FP measurement, performed according to the IFPUG or Nesma process.

The fact that more than half the effort is concentrated in phases 3 and 4 also appears to confirm the reliability of results. In fact, it is popular wisdom that most measurement effort is required by the analysis of data and processes, which is concentrated in phases 3 and 4.

C. An issue with survey one

In Figure 2 it can be observed that the percentage effort data have fairly large variance only for phases 3 and 4. This suggests that some measurers dedicate to phase 3 only the minimum effort that is needed to identify transactions and data files, so that phase 4 requires a substantial amount of work; on the contrary, some measurers probably perform some amount of analysis of transactions and data already in phase 3, so that the effort for phase 3 increases, while the effort for phase 4 decreases, because part of the analysis needed to assign complexity levels to transactions and data files has already been done.

The fact that some activities of phase 4 can be anticipated into phase 3 is critical for our analysis, since we know that most EEM's allow for simplifying or skipping altogether phase 4, while phase 3 cannot be simplified much, in general. As a consequence, the results of survey one can be regarded as non-conclusive. For this reason we performed a second survey, which is described in the following section.

IV. THE SECOND SURVEY

A. The survey

The second survey (whose details are given in Appendix B) was carried out during a meeting of the GUFPI-ISMA (the Italian Function Points User Group – Italian Software Measurement Association), which took place in Rome in December 2016. The questionnaire was first explained to the participants, then the respondents compiled the questionnaire (on paper) and handed it to the author.

The collected answers concerning the relative (percentage) effort for IFPUG Function Point measurement are given in Table VII in Appendix B. Note that respondents were invited to express a specific percentage of the total effort for each phase, rather than a range, as in survey one.

We collected 37 answers. Of the respondents, 36 are certified Function Point Specialist (CFPS) and one is a certified Function Point Practitioner (CFPP). On average, the respondents have over 10 year experience and count over 7000 FP per year. Therefore, the respondents are extremely well qualified, and we can regard their answers as exceptionally reliable in representing current IFPUG measurement practices.

B. The results of the survey

Table II reports the mean relative effort per phase according to respondents. Note that –unlike in the first survey– no normalization was necessary, so we could compute the statistics given in Table II directly from the collected data.

TABLE II. STATISTICS OF PHASE RELATIVE EFFORT

Phase	1	2	3	4	5	6
Mean	16.9	10.2	23.3	31.5	8.1	10.0

Figure 3 shows the distributions of the relative effort dedicated to each phase according to respondents. It can be seen that the opinions concerning the amount of effort that should be dedicated to phases 3 and 4 vary widely. In fact, for both phases 3 and 4 the distance between the minimum and the maximum evaluations is 40% (quite a large variation indeed).

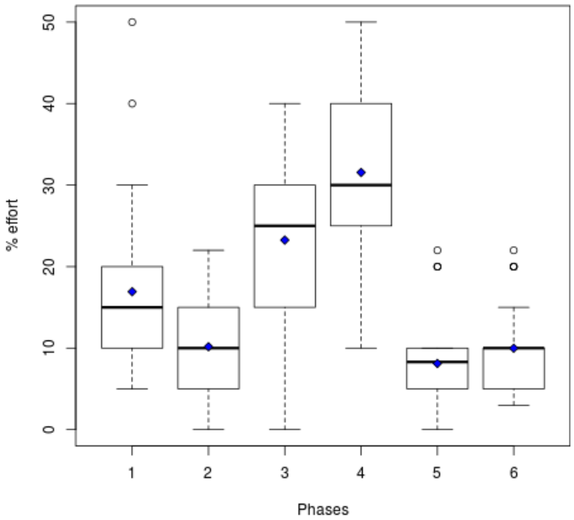

Figure 3. Distribution of relative phase efforts for the second survey dataset.

Figure 3 suggests that there is little agreement on how to perform the FP measurement process. This is confirmed by the visual analysis of raw data (see Table VII in Appendix B). In fact, it seems that some respondents prefer to start measuring after a relatively lightweight analysis of requirements, which results in longer and/or more difficult activities in phases 3 and/or 4: for instance, respondent 9 spends only 5% of the measurement effort in phase 1, thus causing the effort for phases 3 and 4 to increase substantially (in fact, phases 3 and 4 together require 85% of the total measurement effort). On the contrary, some respondents perform a thorough analysis of requirements in phase 1, so that the following phases become simpler and require less work: for instance, respondent 10 spends 50% of the measurement effort in phase 1, then phases 3 and 4 together require only 20% of the total effort.

In practice, it appears that there is not a unique way of implementing the IFPUG measurement process: rather some measurers prefer to collect all the required information in advance, so that the actual measurement phases are facilitated; on the contrary, some other measurers prefer to collect only the minimum information necessary to start; then they explore details when needed, during the measurement phases (phases 3 and 4).

This observation suggested to partition the dataset into two datasets, corresponding to the two approaches described above: in the first dataset we put the data that indicate heavyweight phase 1, in the second dataset we put the data concerning processes characterized by lightweight phase 1.

The data in Table VII suggest to use 15% as the threshold that can be used to partition the data concerning phase 1.

Figure 4 shows the distributions of the relative effort dedicated to measurement phases, when phase 1 is heavyweight (i.e., it consumes more than 15% of the total effort). Table III reports the mean effort per phase when the process is characterized by heavyweight phase 1.

It is easy to see that the mean effort for phases 3 and 4 decreases sensibly (as expected, since most work is done in phase 1). The variability of phase 3 also decreases sensibly, while the variability of phase 4 remains quite large.

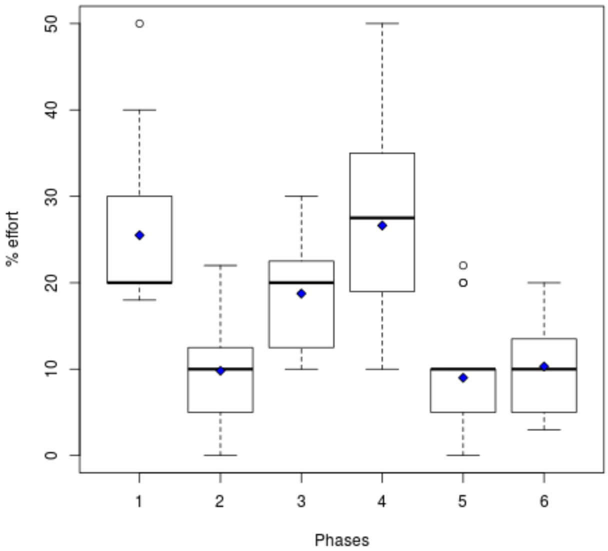

Figure 4. Distribution of phase efforts for the second survey dataset (only "heavyweight" data collection).

TABLE III. MEAN RELATIVE EFFORT PER PHASE, WHEN PAHSE 1 IS 'HEAVYWEIGHT'

Phase	1	2	3	4	5	6
Mean	25.5	9.8	18.8	26.6	9.0	10.3

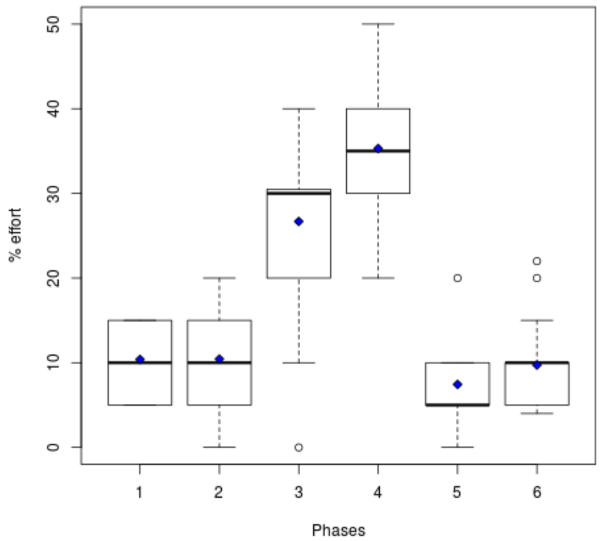

Figure 5. Distribution of phase efforts for the second survey dataset (only "lightweight" data collection)

Figure 5 shows the distributions of the relative phase effort, when phase 1 is lightweight (i.e., it consumes no more than 15% of the total effort). Table IV reports the mean effort per phase when the process is characterized by lightweight phase 1.

It is easy to see that the mean effort for phases 3 and 4 increases (as expected, since little preparatory work is done in phase 1). The variability of phase 3 and 4 decrease, although not much.

TABLE IV. STATISTICS OF PHASE RELATIVE EFFORTS, WHEN PAHSE 1 IS 'HEAVYWEIGHT'

Phase	1	2	3	4	5	6
Mean	10.4	10.4	26.7	35.3	7.4	9.7

V. AN ASSESSMENT OF POSSIBLE SAVINGS

The standard Function Point measurement process according to the IFPUG counting manual can be represented as in Figure 6. Actually, the process in Figure 6 is an abstraction of what really happens in practice, since it ignores loops. For instance, it happens quite frequently that while analyzing a transaction to determine its complexity the need to collect more information concerning the transaction arises, so that the measurer has to go back to the documentation gathering phase.

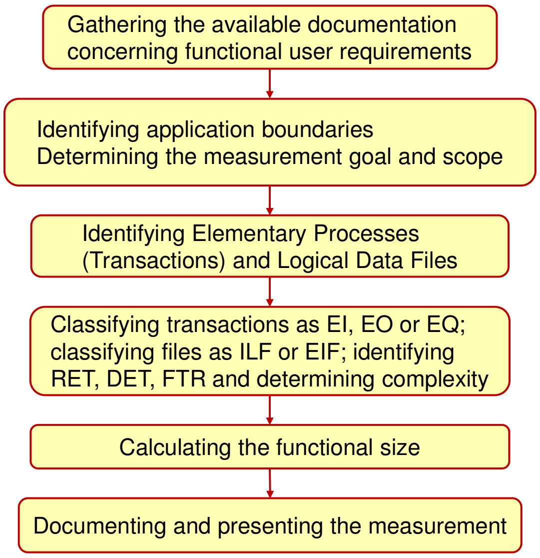

Figure 6. A schematic representation of the FP measurement process.

Now, the most popular EEM's simplify the phase that involves classifying transactions and data files and determining their complexity to a great extent. Among such methods are the NESMA estimated [9][12][13], Early&Quick Function Point [11], simplified Function Point [15] (not to be confused with the Simple Function Point method), and ISBSG average weights (which assigns to each basic functional component the average weight that type of component has in the ISBSG (International Software Benchmarking Standards Group) dataset [14]). For instance, the NESMA estimated method only requires that transactions are classified into EI, EO and EQ, and logical data files are classified into ILF and EIF, but no weighting is required; thus there is little need to analyze the details of transactions and data files, which is definitely the most effort consuming activity in this phase.

When the SiFP method is used, the phase that involves classifying transactions and data files and determining their complexity is skipped altogether, since the SiFP method does neither require that transactions are classified into EI, EO and EQ, nor that data files are classified as ILF or EIF.

Moreover, with both EEM's and SiFP the first and the last two phases are also greatly simplified:

- Since the methods require less details, there is less information to be gathered.
- Computing the size measures is often straightforward with early estimation methods. With SiFP, it amounts to computing formula (1), so there is actually no work to do.
- Documenting the measurement is simpler, since there is less to document (e.g., when there is no notion of complexity, one does not need to document the characteristics of transactions and data that determine their complexity).

According to the findings reported in Sections III and IV. we can (very roughly) estimate the possible savings that can be achieved via EEM's and with the SiFP measurement method.

In Table V, we compute the savings that in principle can be achieved for each phase:

- In column "Est. % savings" we have the percentage of phase effort that can be possibly saved. This is a subjective estimation, based on the considerations reported above.
- In column "1st survey M_i" the average weighted relative effort for each phase (as in Table I) is given.
- In column "2nd survey M_i" the average relative effort for each phase (as in Table II) is given.
- In the rightmost column, the approximate potential saving is computed. Since the mean efforts resulting from the two surveys are different, we often provide a range (form the most pessimistic to the most optimistic hypothesis). So, concerning phase 1, for instance, we can possibly save 50% of the effort, which is between 14.0% (according to the first survey) and 16.9% (according to the second survey). Thus, we have a minimum saving of 50%×14.0%=7% and a maximum saving 50%×16.9%=8.45%: accordingly, we provide an approximate evaluation that the likely saving will be in the 7-8% range.

TABLE V. POTENTIAL SAVINGS WITH EEM'S

Phase	Est. % savings	1st survey M_i	2nd survey M_i	Likely savings (approx.)
1	50%	14.0%	16.9%	7–8 %
2	0%	12.5%	10.2%	0 %
3	0%	35.8%	23.3%	0 %
4	80-100%	23.3%	31.5%	18–30 %
5	90%	6.2%	8.1%	5–7 %
6	50%	8.2%	10.0%	4–5 %
Total	–	100.0%	100.0%	**34–50 %**

Of course, the evaluations given in Table V are based on averages, thus the reader is advised that in specific cases the actual savings could be somewhat different.

Anyway, we have to note that the savings described above are *potential* savings, that is, it largely depends on the specific situations if the possible savings are actually achieved or not. For instance, if an organization represents requirements via UML diagrams, the first phase of the measurement process is greatly eased, since extracting the information required for functional measurement from UML diagrams [24][25] is fairly easy; so phase 1 will require a smaller relative effort, hence savings on phase 1 will not be very large, in absolute terms.

Finally, let us consider the answers provided during the second survey to the question concerning the relative effort required by EEM's. In fact, in the second survey, measurers who use EEM's were invited to directly indicate what is the effort required for measurement when an EEM is used instead of the standard IFPUG process. The frequency of answers is given in Figure 7, where the effort for measurement using EEM's is expressed as the percentage of the effort required using the standard process. Note that Figure 7 illustrates data concerning the spent effort: to derive the amount of savings you have to subtract the percentage effort from 100%: for instance, when the effort spent with EEM's is 40% of the effort required by the standard process, you save 60% of the effort.

The picture shows that most respondents stated that with EEM's the measurement of FP requires 60% of the effort required by the standard IFPUG process (i.e., they save 40%). It is also apparent that we received only 7 answers, therefore we cannot draw general conclusions. Anyway, it is noticeable that these answers are consistent with our computation of the potential savings:

- The most frequent answer indicates 40% savings, and 40% is mid-way in the 34%-50% range of expected potential savings.
- The collected answers indicate saving in the 30%-60% range: they are consistent with our expected potential savings, with just one respondent that is more pessimistic (having indicated that savings amount to 30%) and one respondent that is more optimist (having indicated that savings can amount to 70%).

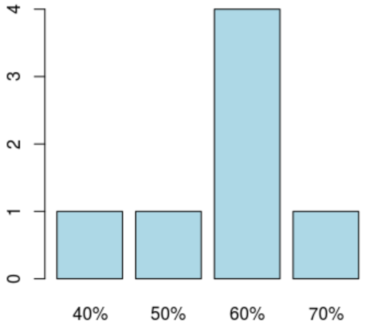

Figure 7. Frequency of answers concerning the relative effort required by EEM's

VI. Threats to Validity

A first threat to the validity of the study is due to the number of datapoints that were collected. Although it was possible to collect less than 40 datapoints in each survey, we strived to guarantee the representativeness of the collected data by eliminating outliers, as well as data that appear incorrect. In any case, the size of the datasets that was finally analyzed is not smaller than many datasets used for empirical software engineering studies.

Concerning the statistical analyses that were performed in this study, they are so simple that it is unlikely that any serious threat to statistically validity actually applies.

Most respondents to the first survey are from Italy, four are form the Netherlands and the remaining ones are from Brazil, Switzerland and Belgium. All the respondents to the second survey are from Italy. The lack of geographic dispersion could be a limit for the generalizability of results. However, most respondents are certified Function Point Specialists (CFPS) or certified Function Point Practitioners (CFPP), thus we can assume that they all follow the process specified in the official manuals [4][5][9][10] (or at least they should be trying to follow such process). Hence, our results should be applicable to all the measurements performed according to the standard counting practices.

VII. Related Work

There is not much literature concerning the cost of functional size measurement. A couple of documents report about the total cost of FP measurement [6][7], but none provides information concerning how the total effort is spread among the various measurement phases.

Some indications are provided by the proposers of EEM's. For instance, it was reported that "*the E&Q size estimation technique has been proved in practice to be quite effective, providing a response within ± 10% of the real size in most real cases, while the savings in time (and costs) can be between 50% and 90% (depending on the comprised aggregation level) with respect to corresponding standard measurement procedures.*" [19]

It was also reported that "*the results found with NESMA estimated fall within a reach of -6% to +15% of the corresponding result found with a NESMA detailed approach, and NESMA estimated FSM is performed 1,5 times as fast as a NESMA detailed FSM.*" [13]

These evaluations are probably optimistic to some extent. However, they are not precise enough to be used for decision making: for instance, it is not clear if the reported savings are evaluated with respect to the whole measurement process or only with respect to the core part (phases 3-5).

VIII. Conclusions

The measurement processes of IFPUG and Nesma FP require a quite detailed analysis of the transactions that the measured software application is expected to provide and of the data it has to manage. Under given circumstances, practitioner may actually consider the standard process excessively expensive or time consuming. To overcome this problem, EEM's have been proposed: via these methods it is possible to obtain faster and cheaper approximate measure estimates.

With similar purposes, but with a different approach, the SiFP method proposes a full-fledged functional size measure that can be used as an alternative to Function Point. The SiFP method –like EEM's– greatly simplifies the functional measurement process.

In any case, managers who have to choose whether to perform a standard IFPUG measurement or an approximate IFPUG FP estimation or a Simple Function Point measurement need to know how much measurement effort can be actually saved.

Since we know which measurement phases are skipped or simplified with EEM's or the SiFP method, to evaluate the possible savings we need to know the relative effort required by the measurement phases that compose the standard IFPUG measurement process. To this end, questionnaires were proposed to professional measurers, in two distinct surveys, and the collected answers were analyzed.

The results of the analysis –given in Sections III and IV illustrate the relative effort required by each phase of the measurement process according to professional measurers.

Considering the overall measurement process, our analysis shows that potential savings are in the 34%-50% range, with respect to the effort required to carry out the standard IFPUG process.

As a final remark, it is interesting to note that both in the first and in the second surveys we found large variations in the effort dedicated to some phases. This indicates that –even though the IFPUG certified measurers are expected to carry out a well-defined standard process– in practice there are a few activities that in some cases are "extended" so to ease the following ones, while in other cases they are performed "minimally," so that more work is demanded to the following phases. Therefore, when evaluating the amount of savings that can be achieved, an organization should first understand what flavor of measurement process they implement, because this determines where the biggest opportunities for savings are located. For instance, when an organization that puts a lot of effort in phase 1 decides to use an EEM, they should reduce the amount of analysis performed in phase 1, otherwise they end up collecting more information than is actually used to get measure estimates via the chosen EEM.

Future work includes:

- Extending the dataset, especially with answers from non-European countries, to make the dataset representative of a larger community of IFPUG users.
- If possible, collecting real effort data from the field, instead of subjective indications provided by measurers. This would make it possible to analyze not only the relative effort, but also the actual effort (in PersonHours, for instance) required by each measurement phase.
- Characterizing the contexts in which measurement is performed, to support the empirical evaluation of the dependency of the relative effort required by measurement phases on the context, and to explore the effects of lightweight or heavyweight activities concerning requirements documentation gathering.

ACKNOWLEDGMENT

The work presented here has been partly supported by the “Fondo di Ricerca d’Ateneo” of the Università degli Studi dell’Insubria.

REFERENCES

[1] L.Lavazza, “An Investigation on the Relative Cost of Function Point Analysis Phases” The 11th Int. Conf. on Software Engineering Advances – ICSEA, 2016.

[2] A. J. Albrecht, “Measuring Application Development Productivity”, Joint SHARE/GUIDE/IBM Application Development Symposium, pp 83–92, 1979.

[3] A. J. Albrecht and J. E. Gaffney, “Software function, lines of code and development effort prediction: a software science validation”, IEEE Trans. on Software Eng., vol. 9, pp. 639–648, 1983.

[4] International Function Point Users Group, “Function Point Counting Practices Manual - Release 4.3.1”, 2010.

[5] IFPUG, “Function Point Counting Practices Manual Release 4.3.1”, 2010.

[6] C. Jones, “A new business model for function point metrics”, http://concepts.gilb.com/dl185, 2008. Last access June 12th, 2016.

[7] “Methods for Software Sizing – How to Decide which Method to Use”, Total Metrics, www.totalmetrics.com/function-point-resources/downloads/R185_Why-use-Function-Points.pdf, August 2007. Last access June 12th, 2016.

[8] L. Santillo, “Easy Function Points – ‘Smart’ Approximation Technique for the IFPUG and COSMIC Methods”, IWSM-MENSURA, pp. 137–142, 2012.

[9] ISO/IEC, ISO/IEC 24750:2005, Software Engineering “NESMA Functional Size Measurement Method, Version 2.1, Definitions and counting guidelines for the application of Function Point Analysis. International Organization for Standardization, Geneva, 2005.

[10] NESMA, “Counting Practice Manual, Version 2.1”, 2004.

[11] “Early & Quick Function Points for IFPUG methods v. 3.1 Reference Manual 1.1”, April 2012.

[12] nesma, Early Function Point Analysis, July 15, 2015, http://nesma.org/freedocs/early-function-point-analysis/ Last access June 12th, 2016.

[13] H. S. van Heeringen, E. W. M. van Gorp, and T. G. Prins, Functional size measurement accuracy versus costs is it really worth it? Software Measurement European Forum, May 2009.

[14] ISBSG, Worldwide Software Development–the Benchmark, International Software Benchmarking Standards Group. http://isbsg.org/software-project-data/

[15] R. Meli and L. Santillo, Function Point Estimation Methods: a Comparative Overview, FESMA conference, pp. 2009.

[16] L. Lavazza and R. Meli. "An Evaluation of Simple Function Point as a Replacement of IFPUG Function Point." IWSM-MENSURA 2014. IEEE, pp. 196–206, 2014.

[17] F. Ferrucci, C. Gravino, and L. Lavazza, “Assessing Simple Function Points for Effort Estimation: an Empirical Study”, 31st ACM Symposium on Applied Computing – SAC 2016, Pisa, April 4-8, pp. 1428–1433, 2016.

[18] Simple Function Point Association, Simple Function Point - Functional Size Measurement Method Reference Manual v01.01, March 2014, http://www.sifpa.org/en/sifp-method/manual.htm. Last access June 12th, 2016.

[19] L. Santillo, M. Conte, and R. Meli. "Early & Quick function point: sizing more with less." 11th IEEE International Symposium on Software Metrics, 2005. IEEE, 2005.

[20] L. Lavazza and G. Liu, An Empirical Evaluation of Simplified Function Point Measurement Processes, Int. Journal on Advances in Software, vol. 6, n. 1-2, pp. 1-13, 2013.

[21] Kwiksurveys site, https://kwiksurveys.com/s.asp?sid=aazttngx1iibno6450647#/ Last accessed June 12th, 2016.

[22] nesma association, www.nesma.org. Last accessed June 12th, 2016.

[23] ResearchGate, http://www.researchgate.net/ Last accessed June 12th, 2016.

[24] L. Lavazza, V. Del Bianco, and C. Garavaglia. "Model-based functional size measurement." Proceedings of the Second ACM-IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2008.

[25] L. Lavazza, and G. Robiolo. "Introducing the evaluation of complexity in functional size measurement: a UML-based approach." Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, 2010.

[26] R. Meli. "Simple function point: a new functional size measurement method fully compliant with IFPUG 4. x." Software Measurement European Forum. 2011.

[27] A. Z. Abualkishik, F. Ferrucci, C. Gravino, L. Lavazza, G. Liu, R. Meli, G. Robiolo. "A Study on the Statistical Convertibility of IFPUG Function Point, COSMIC Function Point and Simple Function Point." Information and Software Technology, 2017.

[28] FiSMA - Finnish Software Measurement Association, FiSMA FSM Method 1.1 (2004).

[29] COSMIC – Common Software Measurement International Consortium, The COSMIC Functional Size Measurement Method – Version 4.0.1 Measurement Manual (The COSMIC Implementation Guide for ISO/IEC 19761: 2003) (2016).

[30] C. Symons, Function point analysis: difficulties and improvements, IEEE Transactions on Software Engineering 14 (1988) 2–11.

APPENDIX A – DETAILS OF THE FIRST SURVEY

A survey about the relative effort required by the phases of Functional Size Measurement

A. The questionnaire

The questionnaire was structured in two sections, described below.

1) About you.

Question	Possible answers
Are you a certified Function Point Specialist (CFPS)?	Yes/No
Are you a certified Function Point Practitioner (CFPP)?	Yes/No
How many years of experience do have in FP counting?	Less than 5 Between 5 and 10 More than 10
How many FP per year do you count on average?	No more than 200 Between 200 and 1000 Between 1000 and 5000 More than 5000

2) Relative effort required by the phases of functional size measurement

According to your experience, what is the relative effort required by the phases of functional size measurement? Please, specify how big is the percentage effort for each phase, according to your experience. Please note that here we consider the measurement performed at the beginning of the project, based on functional user requirements.

Thanks a lot for your answers! If you have any additional comment or remark, or if you want to be informed on the results of the survey, please send an email to: luigi.lavazza@uninsubria.it

Question	Possible answers
Phase 1: gathering the available documentation concerning functional user requirements	0-5%, 6-10%, 11-15%, 16-20%, 21-25%, 26-30%, 31-35%, 36-40%, 41-45%, 46-50%, 51-55%, 56-60%, 61-65%, 66-70%, 71-75%, 76-80%, 81-85%, 86-90%, 91-95%, 96-100%
Phase 2: Identifying application boundaries	0-5%, 6-10%, 11-15%, 16-20%, 21-25%, 26-30%, 31-35%, 36-40%, 41-45%, 46-50%, 51-55%, 56-60%, 61-65%, 66-70%, 71-75%, 76-80%, 81-85%, 86-90%, 91-95%, 96-100%
Phase 3: Determining the measurement goal and scope	0-5%, 6-10%, 11-15%, 16-20%, 21-25%, 26-30%, 31-35%, 36-40%, 41-45%, 46-50%, 51-55%, 56-60%, 61-65%, 66-70%, 71-75%, 76-80%, 81-85%, 86-90%, 91-95%, 96-100%
Phase 4: Identifying Elementary Processes (Transactions) and Logical Data Files	0-5%, 6-10%, 11-15%, 16-20%, 21-25%, 26-30%, 31-35%, 36-40%, 41-45%, 46-50%, 51-55%, 56-60%, 61-65%, 66-70%, 71-75%, 76-80%, 81-85%, 86-90%, 91-95%, 96-100%
Phase 5: Classifying transactions as EI, EO or EQ; classifying files as ILF or EIF; identifying RET, DET, FTR and determining complexity	0-5%, 6-10%, 11-15%, 16-20%, 21-25%, 26-30%, 31-35%, 36-40%, 41-45%, 46-50%, 51-55%, 56-60%, 61-65%, 66-70%, 71-75%, 76-80%, 81-85%, 86-90%, 91-95%, 96-100%
Phase 6: Calculating the functional size	0-5%, 6-10%, 11-15%, 16-20%, 21-25%, 26-30%, 31-35%, 36-40%, 41-45%, 46-50%, 51-55%, 56-60%, 61-65%, 66-70%, 71-75%, 76-80%, 81-85%, 86-90%, 91-95%, 96-100%
Phase 7: Documenting and presenting the measurement	0-5%, 6-10%, 11-15%, 16-20%, 21-25%, 26-30%, 31-35%, 36-40%, 41-45%, 46-50%, 51-55%, 56-60%, 61-65%, 66-70%, 71-75%, 76-80%, 81-85%, 86-90%, 91-95%, 96-100%
Please, specify what measurement method the given data you gave apply to	IFPUG NESMA Other
Please, specify if the given data take into account some type of simplification	No simplification Nesma estimated Nesma indicative Early & Quick FP Other

It should be noted that in the first survey phase 2 was split into two distinct activities: Identifying application boundaries and Determining the measurement goal and scope, which were labeled phase 2 and phase 3, respectively. Anyway, in the paper we have considered phases 2 and 3 together, as we did in the second survey.

B. Answers

Collected answers are in Table VI.

TABLE VI. ANSWERS CONCERNING RELATIVE PHASE COSTS (FIRST SURVEY)

Respondent	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5	Phase 6	Phase 7
1	11-15%	0-5%	0-5%	26-30%	36-40%	0-5%	16-20%
2	16-20%	6-10%	0-5%	36-40%	6-10%	0-5%	16-20%
3	6-10%	0-5%	0-5%	6-10%	46-50%	0-5%	11-15%
4	0-5%	0-5%	0-5%	66-70%	0-5%	0-5%	0-5%
5	0-5%	6-10%	6-10%	36-40%	16-20%	0-5%	0-5%
6	0-5%	0-5%	46-50%	31-35%	0-5%	0-5%	0-5%
7	6-10%	6-10%	0-5%	21-25%	26-30%	0-5%	11-15%
8	26-30%	11-15%	6-10%	11-15%	11-15%	0-5%	11-15%
9	16-20%	0-5%	0-5%	21-25%	21-25%	0-5%	11-15%
10	0-5%	0-5%	0-5%	46-50%	31-35%	0-5%	0-5%
11	31-35%	0-5%	0-5%	21-25%	16-20%	11-15%	0-5%
12	0-5%	0-5%	0-5%	16-20%	0-5%	0-5%	11-15%
13	0-5%	0-5%	0-5%	21-25%	46-50%	0-5%	0-5%
14	0-5%	0-5%	0-5%	41-45%	26-30%	0-5%	0-5%
15	11-15%	6-10%	0-5%	36-40%	11-15%	0-5%	0-5%
16	6-10%	0-5%	0-5%	51-55%	16-20%	0-5%	0-5%
17	6-10%	6-10%	0-5%	26-30%	6-10%	36-40%	6-10%
18	0-5%	0-5%	0-5%	36-40%	36-40%	0-5%	0-5%
19	6-10%	6-10%	0-5%	11-15%	11-15%	0-5%	41-45%
20	31-35%	16-20%	6-10%	26-30%	11-15%	0-5%	0-5%
21	16-20%	6-10%	6-10%	16-20%	11-15%	6-10%	16-20%
22	16-20%	0-5%	0-5%	61-65%	0-5%	0-5%	0-5%
23	0-5%	56-60%	16-20%	66-70%	51-55%	0-5%	11-15%
24	6-10%	6-10%	11-15%	26-30%	11-15%	11-15%	0-5%
25	11-15%	6-10%	0-5%	21-25%	21-25%	11-15%	6-10%
26	6-10%	0-5%	0-5%	41-45%	21-25%	0-5%	6-10%
27	41-45%	6-10%	0-5%	6-10%	6-10%	6-10%	11-15%
28	6-10%	16-20%	6-10%	31-35%	11-15%	0-5%	16-20%
29	11-15%	0-5%	0-5%	66-70%	11-15%	0-5%	0-5%
30	21-25%	0-5%	0-5%	21-25%	21-25%	6-10%	0-5%
31	0-5%	0-5%	0-5%	6-10%	6-10%	0-5%	0-5%

APPENDIX B – DETAILS OF THE SECOND SURVEY

C. *The questionnaire*

The questionnaire is illustrated below.

About yourself

Are you a certified Function Point Specialist (CFPS)?	☐ Yes ☐ No
Are you a certified Function Point Practitioner (CFPP)?	☐ Yes ☐ No
For how many year have you been counting Function Points?	
How many Function Points do you count per year, approximately?	

Did you participate in the survey published on kwiksurveys? ☐Yes ☐No

According to your experience, what is the percentage cost of every phase in the measurement process?
We are considering measures carried out at the beginning of the project, based on Functional Requirements, for new software development projects.

Phase	Description	% cost (using the IFPUG process)	% cost (using an approximate estimation method _________)
1	Collection of documentation concerning functional requirements		
2	Identifying application boundaries and determining the measurement scope		
3	Identifying elementary processes (transactions) and logical files		
4	Classifying transactions into EI, EO, EQ; Classifying data files into ILF and EIF; determining function complexity		
5	Computation of functional size		
6	Documentation and presentation of measures		
Total	**Please check that the sum of phases' costs is 100%!**	**100%**	**100%**

D. Answers

Collected answers are in Table VII.

TABLE VII. ANSWERS CONCERNING RELATIVE PHASE COSTS (SECOND SURVEY)

Respondent	Phase1	Phase2	Phase3	Phase4	Phase5	Phase6
1	15	10	40	20	5	10
2	5	15	30	30	5	15
3	15	10	20	25	20	10
4	15	20	20	35	5	5
5	10	20	0	50	10	10
6	10	10	35	20	5	20
7	5	10	30	40	5	10
8	5	3	15	45	10	22
9	5	0	35	50	0	10
10	50	10	10	10	10	10
11	20	15	20	40	2	3
12	20	5	15	50	5	5
13	20	10	20	20	10	20
14	5	5	40	40	5	5
15	20	0	10	35	20	15
16	40	5	30	20	0	5
17	30	5	30	30	0	5
18	20	20	20	30	5	5
19	20	10	30	25	5	10
20	20	22	10	18	10	20
21	18	15	15	18	22	12
22	15	5	30	35	5	10
23	15	5	30	35	5	10
24	20	10	15	35	10	10
25	30	5	20	30	10	5
26	8.3	8.3	30.5	36.3	8.3	8.3
27	15	8	30	40	3	4
28	15	10	25	35	10	5
29	10	10	10	50	10	10
30	10	15	15	40	10	10
31	30	10	20	20	10	10
32	20	5	25	35	5	10
33	5	20	30	30	10	5
34	10	20	30	30	5	5
35	30	10	10	10	20	20
36	10	10	30	30	10	10
37	15	5	35	25	10	10