The African origin of public policy, programme and project monitoring and evaluation (M&E) practices

Dominique E. Uwizeyimana*©✉

*Director and Professor in the School of Public Management, Governance and Public Policy, University of Johannesburg, South Africa.

Abstract The literature has robust contestation about the origin of public policy, programme and project monitoring and evaluation (M&E). Cloete (2016) argues that public policy, programme and project M&E originated in Europe and America and was introduced in Africa by colonisers who divided the African continent among themselves in the early 1880s. Cloete (2016) does not explain why there was no public policy, programme or project M&E in Africa or any other non-western part of the world before the arrival of western colonisers. Mouton (2010) also argues that M&E was introduced in Africa by international development organisations and financial organisations in the 1980s. According to Mouton (2010), these western development organisations and financial institutions introduced M&E to the African continent as a condition for accessing foreign aid and as a funding requirement for the projects these institutions were funding. Neither Mouton (2010) nor Cloete (2016) accounts for ample literature which demonstrates that M&E practices existed in Africa (and in China) thousands of years before the arrival of the international development organisation and international financial organisations in these parts of the world. This paper uses the Secondary Data Analysis (Archival Study) approach to provide a systematic and chronological analysis of the available literature to trace the genesis of public policy, programme and project M&E as a practice. The findings in this paper indicate that M&E could potentially have become a practice in African countries such as Egypt and Asian countries such as China long before it was a practice in the U.S. and European countries. However, evidence suggests that academics and practitioners in the U.S. and European countries could have compiled the first recorded theoretical academic content about M&E models and approaches currently being used to train people who want to take up a career or profession in M&E.

Keywords: M&E, public policy, programme evaluation, project, evaluation

1. Introduction

It is often argued that to understand an academic topic's origin or historical background, one must trace its research and publication history (Auriacombe 2013). Therefore, according to Auriacombe (2013), one must investigate the historical background of evaluation research to understand better what public policy, program, and project evaluation are all about. There are some essential truths to Auriacombe's (2013) argument. Still, this author does not account for the fact that some ancient civilizations did not write down their history. Instead, they used oral history to pass the information from one generation to another. Researching the historical background of evaluation practices (or as a discipline and field of study) in Ancient Africa is difficult because knowledge was mainly transferred from one generation to the other through oral history. Therefore, while developed countries’ history is well kept in books, articles, etc., the fact that most literature is from Western countries risks giving us a skewed account, heavily biased toward Western countries. The difficulty in finding written records about how M&E was practiced in Ancient Africa is further complicated because colonizers have stolen these records and have hidden them in different museums in their countries. This paper will show that the Palermo Stone, the only credible ancient description of M&E practice in ancient Egypt, is kept in Palermo, Sicily, Italy. If there are written records about M&E practices in ancient Africa, there are safely hidden in private home libraries and are inaccessible to modern researchers.

The following paragraphs discuss the different stages in the history of public policy, program, and project M&E as a practice to determine its origin or genesis, as promised in the opening paragraphs of this paper. But, before doing so, the paper starts with the conceptual analysis of concepts related to the topic of the research and will provide a detailed analysis of the following:

2. Conceptual and contextual analysis

One must first establish a working definition for the essential concepts to study the genesis of public policy, programmes and projects monitoring and evaluation (M&E) practices. Basheka and Byamugisha (2015) state that the term practice implies that the field or an area of work attracts enough people to work in it on a regular basis. As Basheka and
Byamugisha (2015) continue to argue, the fact that M&E practitioners are currently found in many government departments, civil society organisations (CSOs), private and public consultancy organisations, and academia means that M&E is now a fully-fledged practice.

For the analysis in this paper, the concept "Genesis" comes from Ancient Greek "γένεσις" or genesis", which means "the origin or coming into being of something, source, beginning, nativity, generation, production," meaning "origin", "in the beginning" (Merriam-webster. 2022). The English concept "Genesis" is like a Latin word derived from the Greek "gignesthai" which means "to be born," and can refer to the origin or beginning of anything from a heavenly body to an idea" (Merriam-webster 2022). Therefore, while the focus of this paper is on M&E as a practice [due to the limited space in this paper and time], the genesis of public policy, programme and project M&E means its "originem", i.e., origin/s, root/s, foundation/s beginning/s, source/s, where it rises, the place where it began, where it sprang into being (Vocabulary.com 2022) as a practice.

The other central concept in this paper is monitoring (M). Monitoring is the routine tracking and regular collection of data about programs and projects’ activities (USAID 2022). However, while monitoring is integral to the evaluation, the two concepts are different. According to The American Evaluation Association (AEA 2017), evaluation (E) involves using the information and data collected through monitoring to do proper SWOT analysis (strengths, weaknesses, opportunities, and threats) of policies, programs, and projects being implemented; as well as the people, products, and organisations that are implementing them to improve their efficiency and effectiveness. Evaluation can lead to continuing interventions if they are on the right track to success. Introducing improvements to the initiatives if they can succeed with little adjustments or corrections or discarding them if they have no chance and trying to salvage them is a waste of time and resources.

Kibeyi (2019) argues that M&E are integral part of the policy, programme and project cycles from conception, through to planning, implementation, termination, and beyond. In contrast, evaluation is "a systematic and intentional process to measure the progress, short-term and long-term results (output and outcomes), and impact of a project" monitoring focuses on capturing and storing data regarding activities and progress daily. In other words, evaluation goes beyond just checking the project’s timely completion within the budget Uwizeyimana (2020). When evaluation is conducted before the implementation starts it is called diagnostic evaluation, when it is conducted during the implementation or execution process, it is called formative evaluation, and when it is conducted after the implementation has been completed it is called summative evaluation (Bakewell et al 2008). Irrespective of when evaluation is performed, its objective is always to improve current and future management decisions. To assess the extent of change and the causal link between the status quo before the implementation starts and the observed change at the time evaluation is conducted, M&E may produce one or a combination of the following comparisons:

a) Comparisons of data collected over time or time series analysis (i.e., tracking one or more indicators over a long time to see how they change. This type of data usually comes from the service providers' records. Still, it may also come from "focus group discussions and community surveys, random sample surveys of intended beneficiaries or other techniques" (Kibeyi 2019).

b) Comparisons over space/areas (i.e., two or more different areas) called cross-sectional analysis. Making comparisons over geographical space or areas requires comparing one population group with another (Kibeyi 2019). This type of evaluation may include comparing one area that has received the intervention with another that did not receive the same intervention. The second area is used for control (with no intervention) to determine the difference between the two areas (i.e., the real impact of the intervention).

c) Counterfactual comparisons which answer questions like "what would have happened to this target community if no project or program intervention was implemented? In addition, it also answers questions like what could have happened if the programme or project had been designed and implemented differently in the same area (Kibeyi 2019).

3. Theoretical analysis

The purpose of public policy, programme and project M&E is to determine a change (i.e., output, outcome and impact) and the causal link between the observed change and the implemented policy, programme and project intervention. The theory of change (ToC) is often used to explain these changes and causal relationships. A ToC represents the normative, structural and operational mental models underlying the success or failure of government’s intervention (Uwizeyimana 2020). According to the USAID (Uwizeyimana 2020), a well-formulated ToC must achieve the following objectives:

- To identify all the internal and external conditions necessary and sufficient to achieve a desired goal (i.e., desired change or output, outcomes or impact) starting from the situation before the implementation as the baseline.
- Clarifying the causal relationship i.e., how and why the interventions are expected to cause the desired changes.
- Explaining the sequence in which the events leading to change should occur in order to cause the desired changes.
- Answering the “why” questions (i.e., the rationales) behind this sequence of changes (Woodrow and Oatley 2013).

https://www.malque.pub/ojs/index.php/msj
- Identifying other public and private interventions and conditions (both internal and external, local and international) that must happen and cause the desired changes.
- To recognise where there might be gaps in the environmental analysis and how to close them or minimise their effect.
- To know which indicators to monitor when monitoring progress.

In summary, Uwizeyimana (2019) and Uwizeyimana (2021) argue that a ToC explains why following a particular course of action will lead to the desired change in the prevailing context. This can be summarised in the following (Woodrow and Oatley 2013) statement:

"If we do X actions..., then Y (output, outcomes, and impact) will happen because Z (conditions)... are in place" (Woodrow and Oatley 2013), in Uwizeyimana (2019) and Uwizeyimana (2021).

According to Masuku and Ijeoma (2015), “modernist scholars argued that “the evolution M&E started in the western countries”. However, many scholars such as Shadish (Cook and Leviton 1991; Kusek and Rist 2008; Stockman 2011; UNESCO 2016; Matsiliza 2019) tend to disagree and argue that “M&E is not a new concept in Africa”. and They argue that it has “existed in the form of traditional governance during the Stone Age” (Masuku and Ijeoma 2015). As Masuku and Ijeoma (2015) put it “it is so unfortunate that the current literature on M&E has, for a very long time, “overlooked the evolution from the Stone Age context”.

The following paragraphs trace the different stages in the development of M&E as a practice to determine its origin or genesis.

4. M&E practices in the animal kingdoms - Since the beginning of life on Earth

Literature suggests that the practice of evaluation was naturally done by animals such as apes (chimpanzees, chimps, monkeys, orangutangs etc.) measuring the depth of a river or a pond to find out how deep it is before crossing the river or diving. Evaluation as an animal's instinct survival mechanism to avoid drowning was behind this practice in the animal kingdoms. The following picture (Figure 1) shows "Mego the Orangutan" using "a long, staff-like stick to measure the depth of the river to find out whether the river was too deep to cross. (Finnis 2014) argues that "Mego the Orangutan was rescued by the Nyaru Menteng Rescue and Rehabilitation Centre in Central Kalimantan, Indonesia. A few days after the Orangutan was released into the wild in 2014, saw him at the bank of the Joloi River holding a long branch of a tree (Finnis 2014). According to the eyewitness in the park, as stated by Finnis (2014) the Orangutan slowly inserted it into the water to measure the depth and only made his way in, after realising it was safe to cross”. But unluckily, things did not go well with Mego as he latter lost grip of the branch and hopelessly watched it float away. After realising that it would be dangerous to continue the journey across the river without it, the clever animal decided to turn back to the shore of the river (Finnis 2014).

It can be argued that Mego the Orangutan's act of using a stick to measure the depth of a river is almost equivalent to collecting the data and/or information, its decision to wade in the river after concluding that it was not too deep, and its ability to decide to turn back to safety after losing the branch (i.e., a measuring instrument) because it deemed it to be dangerous to cross the river, without it meets the basic and minimum analytical requirements of modern M&E practices.

Evaluation implies the ability to systematically analyse the information and make appropriate judgement or decisions, or recommendations to yourself and others (Uwizeyimana 2020). Evaluation requires the evaluator to make decisions or choices, to “estimate, judge, defend, criticise, and justify” (Bloom et al 1956; Huitt 2011).

The fact that animal practiced evaluation as a natural survival instinct and the fact that "men" evolved from the apes suggest that evaluation could have been practiced since the beginning of people perhaps also as an instinct for human survival. The apes, and especially orangutans share 98 per cent of human’s DNA, explains “their remarkable similarity to us” and why they are adaptable and innovative like humans (McEntee, in Finnis 2014).

Figure 1 Mego the Orangutan testing the waters to checks the depth of the river with a stick before crossing. Source: https://www.dailymail.co.uk/news/article-2746844/Don-t-depth-Incredible-pictures-orangutan-using-stick-check-river-safe-cross.html. Accessed on: 21 October 2022.
5. M&E in Ancient human kingdoms and societies

Appraising or examining a phenomenon's relative significance is a natural and intrinsic part of every human being. Accordingly, (Shadish 1991) use the example of someone in the pre-historic era trying to “roast meat to see if it can taste better than raw meat. Knowingly or unknowingly, such ancestors of ours were conducting evaluation and their findings of such evaluation processes are still beneficial to people in modern societies across the globe. In addition, “someone who, in the pre-historic era, ventured into discovering “whether certain fungi, leaves, roots, plants, insects, birds, animals, and sea and air creatures, were edible or not, or if the work could be carried out by using one particular type of tool than with another, that person was conducting evaluation (Stockman 2011). Hence the history of human existence is filled with difficulties and challenges they must overcome to survive, many forms of evaluation practices were conducted by human ancestors in the pre-historic era. The time for such discoveries is not recorded anywhere in the history books because humans in the pre-historic era had not started documenting their history. Still, no one can deny that these were life-changing evaluations and that such evaluation practices have had and continue to have real effects on the lives of people living today and those who will live in the future. Therefore, one has to look at the times and places where human beings lived first in the pre-historic era to determine the origin of evaluation practices (Stockman 2011).

The most available scientific discovery shows that the human evolution stared as early as 300,000 years ago as demonstrated by “The 300,000-year-old bones and stone tools” of Homo sapiens ever found" (Yong 2017). The earliest fossilised remains of Homo sapiens (or primate species) include bones and the stone tools which were recently discovered in a place called Jebel Irhoud "around 62 miles west of Marrakesh, the capital city of Morocco" (Yong 2017). These remains are believed to belong to the “Homo sapiens, members of a nascent ape species that had spread across Africa” around 315,000 years ago (Yong 2017). Before the Moroccan discovery, the oldest known fossilised remains of Homo sapiens were the two fossilis discovered in Ethiopia which are estimated to be between 160,000 and 195,000 years old (Yong 2017). The two discoveries of the earliest Homo sapiens ever found confirms modern scientific research which shows that Africa is the "Cradle of Humankind" (Yong 2017). Hence the identification of the problem, and alternative options to solve them, is dated back to the early stage of human evolution, and the earliest fossilised remains of Homo sapiens were recently discovered in two African countries (Morroco and Ethiopia), the one can safely argue that the earliest M&E practices were initiated by Ancient African communities.

5.1. Government monitoring and evaluation in ancient times

According to UNESCO (2016), public policy, programme and project monitoring practices have been an integral part of the education process in most developed and developing countries over the last few decades. As UNESCO (2016) states "practices, such as school attendance registers have been part and parcel of the schools’ performance management strategies for many years. Furthermore, "basic education data, such as the number of schools in a particula geographic location, the number of students enrolled in schools and the number of teachers in a particular geographical area or school have been part of national education statistics for many years (UNESCO 2016). In fact, all developed countries, such as the U.S. and other western Europe countries, have well-established M&E systems. According to UNESCO (2016), even less developed countries, such as Bangladesh, Sri Lanka, Kenya, Uganda and Tanzania, have had simple public policy, programme and project monitoring tools since early 1990.

The above argument support UNESCO's argument that governments monitoring and evaluating their performance is not new Kusek and Rist (2008). The thought and influence of M&E were recorded in ancient human kingdoms about 5000 years ago. This is "when ancient governments and officials would take note of the performance of their livestock and grain sales, the quality as well as the balancing of the financial books" (Matsiliza 2019). According to Matsiliza (2019), although in those earlier days, M&E might not have been “as prominent and advanced as it is today, the modern principles of ensuring quality and assessing results are much applicable in modern-day governance and policy-making”’. This M&E process which "assisted in planning ahead for sales and satisfaction and means of improvement in future instances" in ancient kingdoms, is "the same code which is being followed for present-day M&E" (Matsiliza 2019). Matsiliza (2019) argument confirms Kusek & Rist (2008) argument, which acknowledges that governments' M&E of their own performance is not new. Literature suggests that monitoring and evaluation were regularly practised when the Egyptian Pharos built the Pyramid of Giza. The Great Pyramids of Egypt were constructed in about 2630 BC (Meredith 2014). That was 1630 years before Stonehenge (erected in England estimated 2000 B.C.) and 2000 years before Mexican Sun Pyramid built1500 B.C.), and over 5000 BC before the first western travellers (led by Vasco Da Gama in1498) arrived in Africa (Meredith 2014). There was a specialist in performing the important task (i.e., M&E) to ensure that the Pyramids met certain specific structural and engineering criteria and their alignment with celestial planets and stars, but also to ensure the completion of the work on time. The scientific details care taken by the people who built the Pyramids of Giza etc., and the astronomic characteristics suggest that these people were far advanced in all scientific fields by about 3000 BC. This is, for example, the view held by (Haughey 2014), who argues that
"the history of M&E can be traced back to Egypt in 2570 B.C." and that it was practiced "during the construction of the Great Pyramid of Giza". As Haughey (2014) puts it:

"The Egyptians built the pPyramid'shistorical records, which portray that they were managers for each pyramid cycle to oversee their work to completion. Some degree of planning, execution and control was adopted to manage this project". Therefore, even though the term M&E was not coined at that time, "there was a monitoring and evaluation level in this". Therefore, as Haughey (2014) concludes, "based on this historical evidence", M&E of projects and project management are not new concepts.

There is a consensus among authors such as (Kusek and Rist 2008; Clayton 2012, Meredith 2014; Birasa 2019) that "The Pharaohs built the pyramids with exceptional precision, and that such work would not have been possible without exceptionally skills managers dedicated to overseeing (i.e., monitoring and evaluation) the work for each face of the Great Pyramid. The fact that the pyramids were completed in a specific time and according to the plan suggests there were managers and leaders behind who were monitoring and evaluating the progress of the building projects when the Pharaohs built the great pyramids of Giza in 2750BC (Haughey 2014).

5.2. The Palermo Stone and the Nilometer as the oldest M&E writings and instruments in Ancient Egypt

One of the oldest written pieces of evidence to demonstrate that the Ancient Egyptians were far advanced in terms of M&E during the Early Dynastic Period the Ancient Egyptians is the "Palermo Stone". Schäfer (1902) states that "the Palermo stone are a "set of several pieces of a large black basalt slab", which are “written in hieroglyphics (i.e., an Ancient pictorial writing) on both sides"). More importantly, regarding the discussion in this article, the Palermo Stones contain the information about how the Nile River was monitored in Ancient Egypt (from around 3200 to 2350 BC) (Schirò 2021). During the 700 years period, "the height of the Nile was recorded as 6 cubits and 1 palm (about 3.217 m or 10 ft 6.7 in)” (Clagett 1999). As correctly indicated by Schäfer 1902:

Written in Ancient Egyptian Hieroglyphs, "the Palermo stone has short, year-by-year accounts of major events that took place during the reign of different pharaohs. These events are mostly ceremonies, festivals and rituals" and are written in the shape of a palm branch. The bottom of each register records the Nile heights. In Schäfer's (1902) own words:

The Nile heights are in ancient Egyptian units of lengths:

<table>
<thead>
<tr>
<th>Unit</th>
<th>Abbreviation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digit</td>
<td>dbt</td>
<td>0.74</td>
</tr>
<tr>
<td>Zep</td>
<td>szp</td>
<td>4 fingers</td>
</tr>
<tr>
<td>Moh</td>
<td>mh</td>
<td>7 palms</td>
</tr>
</tbody>
</table>

The strokes stand for multiples of the signs. For example, 𓊱𓊱𓊱𓊱𓊱 2 dbt 2, means 2 cubits and 2 fingers. This converts to 2 × 7 × 4 + 2 fingers = 58 fingers = 42.92 inches.


The following is an example of the Palermo Stone (Figure 2) discussed in the above section.

Figure 2 The Palermo Stone (Palermo Museo Archeologico. (Source: http://xoomer.virgilio.it/francescoraf/hesyra/palermo.jpg; and also http://www.catchpenny.org/thoth/Palermo/index.htm (Accessed on: June 1, 2023).
In addition to the Palermo Stone, there is evidence that the Nilometer was the oldest known M&E in Ancient Egypt.

6. The Nilometer is the oldest known M&E in Ancient Egypt (>5200 ago)

There is overwhelming evidence to demonstrate that the Ancient Egyptians used one of the very sophisticated scientific measuring instruments known as the Nilometer to monitor, record and evaluate the Nile River's seasonal variations and their impact on flood and harvest (Kiprop 2018). Archeologic evidence suggests that Nilometers originated in Pharaonic times (i.e., from 3200 BC until 332 BC when the country fell under the Macedonian conquerors). The fact that the Nilometer was used to measure water levels as early as 5200 years ago suggests that nilometers are the oldest known monitoring and evaluation instruments (Figure 3).

The analysis of this information helped the people in charge of monitoring and evaluating the Nile River. The priests’ and pharaohs’ (and later, Roman or Arab leaders) ability to record data (daily) and to analyse the data to predict whether the harvest was going to be good or whether there was going to be famine (Mortel 2022) is like current monitoring and evaluation practices in areas of public policies, programmes and project implementation. The skill of predicting nature (quality and levels of the) floods also played a significant administrative and political role in Ancient Egypt civilisations because the quality of the Nile River floods determined the levels of tax paid by citizens in the region in that given year (Mortel 2022). For example, the average inundation suggested an excellent harvest and therefore more taxes to be collected during that annual agricultural cycle, a below average inundation would result in famine and less tax collection. An unusually above-average inundation was catastrophic as it meant that most cultivated land and crops would be washed away (Mortel 2022). The above-average inundation would usually lead to massive destruction of a large portion of the nation's infrastructure constructed along the floodplain and less tax collection. Trend analysis of historical records about the Nile River flooding cycles indicates that "one out of four years on average (or about 28%) between 622 and 999 CE, experienced flooding that did not meet the people's expectations over the same period of about 102 of the years" (Fagan 2010).

![Figure 3 The Nilometer used by Ancient Egyptians to monitor the Nile River flood level (Source: Mortel 2022).](https://www.malque.pub/ojs/index.php/msj)

In fact, ancient Egypt was not the only non-western country with advanced M&E practices in the pre-historic era. Several authors such as (Mouton, 2010; Shadish and Luellen 2011; Alkin and King 2016; Kabeyi, 2019) have traced the origins of M&E back to 2200 B.C. in China, where evaluations were conducted on employees. Shadish et al (1991) assert that the idea of planned social evaluation may be traced back to personnel selection in China around 2200 B.C. In addition, "China started initiatives like program assessment in the public health and education sectors in the eighteenth century" (Alkin 2012).

7. Literature is very scant on the M&E practices between 2200 BC and 202 BC

However, the Great Wall construction in China in 202 B.C. marked another historical event that influenced project management. The Qin Dynasty or Ch'in dynasty, which gave its name to China (founded by Qin Shi Huang, who became the First Emperor of Qin and lasted from 221 to 206 BC), is credited for establishing a system in which the labour force was organised into three major groups: criminals, soldiers and ordinary people. The Qin dynasty commanded these groups. This system, established by the Qin dynasty, led to an emergency of organising, controlling and centralising command structure in the project execution and monitoring and evaluation practices. The centralised command and control structure in these ancient Chinese times ensured proper work planning was done (i.e., executed) and completed on time. Ensuring that projects met their objectives when completed required adequate monitoring and evaluation of these projects.

In addition, there is not enough literature on what happened between the 202BC and the 15th Century. However, a close analysis of the literature shows that some African indigenous people had practices which included some M&E practices like what is done today. Such examples include the tradition of Imihigo in ancient Rwanda. It is unclear when Imihigo practices were introduced in Rwanda, but such practices can be traced back to the 15th Century. (Scher 2010) asserts that Rwanda’s Imihigo is rooted in the country’s pre-colonial cultural practice. During the feudal system era, especially during king
Mutara Rwabugiri reign (approximately 1853/1860-1895), warriors would publicly vow kill the enemy of the king or conquering neighbouring kingdoms on behalf of the king or doing some other acts deemed to require exceptional courage. Once the person who made the vow delivered his promise, he was given a reward (Kugorogerwa), which ranged from cows, the king's daughter for marriage, being made ruler of a region or a combination of these rewards. Once he failed to deliver on the vow, he was either killed or humiliated in a public ceremony.

The practice of Imihigo since the 15th Century confirms Potter and Kruger (2001) argument that the activities resembling modern/current projects, programme and policy evaluation had existed for many centuries before the advent of "modern" programme and policy evaluation in the 1960s. The Imihigo practices, which were part of Rwanda's tradition before the arrival of Germans (1895-1919), and Belgian colonisers (1923-1961) (Kayibanda 1962, Uwizeyimana 2020), show that M&E was part of Rwanda's cultural practices and cultures. The post-1994 Rwandan government re-launched the course of Imihigo in 2006 (Scher 2010). Modern Performance Contracts (imihigos) involves the signing of performance contracts between the President Paul Kagame and leaders at district and local government levels of government (Murasi and Auriacombe 2022).

According to Klingebiel (2019), the signing of the Imihigos between the President and mayors at the district level act as a policy management tool designed to ensure the implementation of national policies at the district level. As in the olden days of traditional Rwanda, current leaders who succeed in achieving their imihigo targets are publicly rewarded, while the ones who fail to deliver on their imihigo face public humiliation and may get dismissed from their jobs Casley and Kumar (1987).

8. Public policy, programmes, and projects M&E practices from 1700 to date

According to Rossi and Freeman (1993) and Babbie and Mouton (2001), policy, programme, and project evaluation-like activities although different from Results-Based M&E, Kusek and Rist (2008) were evident in the fields of education and public health in the 18th Century. However, the following are just a few examples demonstrating that evaluation practices have existed in almost all government endeavours and have gone through different stages of development and perfection to become what they are today since the 1790s.

<table>
<thead>
<tr>
<th>Era</th>
<th>Age</th>
<th>Description</th>
</tr>
</thead>
</table>
| 1.  | Age of Reform/Measurement (1792-1900). | The Age of reform (1792-1900), also known as the "Age of Measurement", is said to be the first generation of modern evaluation. During this first generation of contemporary evaluation, evaluation mainly focused on "measurement" and favoured statistics and surveys. This era was also characterized by performance measurement in schools (Beywi and Gullickson 2022). The U.K. seems to have been at the forefront of educational evaluation practices and methodological development during this era. Literature suggests that it is in that quantitative marks were first used in Britain to assess the performance of students' assignments in 1792 (Hogan 2007). Professor William Farish (1759–1837), is credited for having been the first person to use quantitative mark allocation to exam questions so that he can measure individual university students' examination performance in 1792. However, while Prof William Farish's use of written examination and quantitative marking is said to be the earliest known official usage of evaluation in education, this does not mean that students' assessment was first done in the U.K. Fahim and Zoair (2016) analysis of the "Education systems and practices in Ancient Egypt before the end of the Graeco-Roman period" shows that practical exams, written exams, and oral exams were used the main assessment methods to determine whether the students have acquired the necessary skills in the different educational fields and subjects. According to Fahim and Zoair (2016) analysis of the education level, its objectives, teaching, and learning, as well as its assessment methods during the Graeco-Roman era (i.e., 332 BC-395 AD) (i.e., 100 years period after the Persian rule over Egypt, 395 AD-495AD); shows that systematic teaching and systematic evaluation or assessment of student's academic performance was being practised. According to Fahim and Zoair (2016), during the Graeco-Roman Period, all Egyptian children, including immigrants such as the Jewish Moses in the Bible, had to gain different kinds of scientific knowledge to be considered well educated and ready to perform their societal roles. In fact, as Fahim and Zoair (2016) put it, "The Hellenistic education systems and processes, inherited by current generations from the ancient Egyptian and Athenian models, has kept its manners with less to no fundamental changes. In fact, it is often argued that the only small modifications which occurred in teaching and assessment methods occurred during the Roman Period. Otherwise, the influence of the Egyptian educational model was more than that of the Roman model. If one takes into account Fahim and Zoair (2016) argument, it becomes clear that the use of quantitative marks introduced in 1792 by William Farish (1759–1837) to assess the performance of British students' assignments (Hogan 2007) was not a new assessment method, but most likely a revised assessment method of the methods that might have been in use in Ancient Egypt many millennia before William Farish's (1759–1837) said invention. Therefore, while the Age of Reform (1792-1900) or "Age of Measurement" evaluation was primarily
developed and practiced in the context of education in the U.K. and probably in the U.S. in the 18th Century (Beywi and Gullickson 2022), evaluation practices during this era were heavily inspired by other disciplines (such as agronomy which used "rigorous experiments in plant cultivation" (Rossi 2004), and the monitoring and evaluation of the Nile River flooding seasons which have been practised many millennia in Ancient Egypt as already mentioned in previous sections of this article.

2 19th Century: The 19th Century seems to have been dominated by two countries, namely the US and the UK. In these two countries M&E evaluation studies were undertaken commissions appointed by governments to measure the success or failure of government’s initiatives in critical areas such as the educational, law and health sectors. At that time, the U.S. government appointed many "Presidential Commissions" to conduct large-scale evaluations, examine evidence and judge various programmes, including agriculture, to which agricultural practice yielded the most significant crops. The studies conducted in the early 1900s by the American government are often regarded as the first government-driven evaluation study (Chellimsky 2006). In 1815, the U.S. by the Army Ordnance Development developed evaluation regulations, which are recorded as the first formal evaluation activities (Rossie et al 2004). School performance evaluation was popularised in 1845, when it was formalised for the first time at Boston University, U.S. (Rossie et al 2004).

3 1917: The Gantt chart by Henry Gantt (1861-1919): 1917 was characterised by development of project management tools such as the Gantt Chart, which facilitated M&E practices. Created by Henry Gantt (1861-1919), the Gantt chart is widely used in project planning, implementation management and evaluation (Haughey, 2014). M&E practices benefited from the development of the Gantt Chart as it became a tool for monitoring the progress of programmes and project activities (Kabeyi 2019).

4 Age of Efficiency and Testing (1900–1930): During the 1900–1930 period, scientific management was centered on observation, quantitative measurement and analysis, and efficiency and effectiveness. The 1930s also saw a significant change in the public administration sphere when the American government allocated between $40 and $50 million for evaluation research in Agriculture, education etc., by the American government.

5 Tylerian Age (1930-1945): The Tylerian Age (1930-1945) era, or the post-World War II period (1939-1945), was the catalyst for spreading western evaluation practices to developing countries, especially those that depended on foreign aid. Western donor agencies and global financial agencies especially the IMF and the World Bank consider evaluation an essential programme management tool. M&E became conditional for providing financial (and material) aid to aid-dependent countries. As a result, evaluation "began to emerge both at the country level and in the U.N. system in the early 1950s (Mouton 2010). This period is called the Tylerian Age (1930-1945) because Ralph W. Tyler (1902-1994), and his colleagues championed the importance of evaluating government programmes to find out whether they have accomplished their stated objectives (Mouton 2010). The work of Ralph Tyler and his colleagues was "the catalyst in establishing evaluation as a "distinct field" (Potter and Kruger 2001). The need for monitoring and evaluation of public policies and programmes resulted from the Great Depression of the 1930s, which caused many western governments to increase their budget spending on social and economic programmes to assist with the recovery of their economies. There was a need for objective-based tests to ensure the effectiveness of these government social programmes in solving the socio-economic problems facing many western citizens during and after the Great Depression of the 1930s. Ralph W. Tyler’s (1994) influence was mostly felt in the field of testing and measuring the performance of education programmes and school curriculum. He is credited for transforming the idea of testing and measurement into a global concept that is now called evaluation (Education Encyclopaedia 2022). The (1930-1945) was also characterised by Longitudinal systematic study (education, U.S.): For example, in 1941 Ralph Tyler presented the result of an eight-year investigation of an educational objective and the measurement of their impact. Some consider Ralph Tyler to be the catalyst of a process that led to the establishment of evaluation as a distinct field or practice and study (Alkin and Christie 2004).

6 The Age of Innocence (1946-1957) During the Age of Innocence (1946-1957) period, Americans mentally moved past World War II and the Great Depression (1930s) and started making monitoring and evaluation part of the government policies and programmes. During the Period (1956/7), M&E greatly benefited from the development of the Critical Path Method (CPM).

9. Conclusion and recommendations: what does the above historical review tell us about the origin of M&E as a practice?

The objective of this paper was to use the Secondary Data Analysis (Archival Study) approach to provide a systematic and chronological analysis of the available literature to trace the genesis of public policy, programme and project M&E practices. The study suggests that M&E was practised many years ago in the animal and ancient human kingdoms. The findings also indicate that "to appraise, examine or simply evaluate the relative significance of a phenomenon is part of the human natural and intrinsic phenomena. This finding suggests that M&E is as old as humankind. Thus, one must find out where the earliest human beings have been found to know the origin of M&E practices. Archaeological and historical records show that the earliest Homo sapiens ever seen are on the African continent. This suggests that the oldest M&E practices used by humans, such as roasting meat to find out whether it tests better, or the use of a stick by human ancestors, such as Homo Sapiens, to measure and assess the depth of a river to avoid drowning could have originated from Africa. This finding is also
supported by the literature, which shows that M&E practices, such as systematic evaluation of crops, have been recorded 
bout 5000 BC in ancient Egypt. This is the time Ancient Egyptians built the Pyramids of Giza. Old records such as the Palermo 
Stone show that the Pharaohs who built the pyramids kept records, especially on the Parlemo Stones, showing that there 
were managers for each face of the Great Pyramid. These managers were responsible for overseeing (i.e., monitoring and 
evaluating) the work to completion. While the literature reviewed also show that M&E was commonly practised in Ancient 
China as social evaluation with government staff selection in China early as 2200 B.C.; the fact that the Palermo Stones 
contain information about how the Nile River was monitored in Ancient Egypt (from around 3200 to 2350 BC, approximately 
(>5200 ago), is further proof that M&E practices originated from the Ancient African people who inhabited the northern part 
of Africa during this era.

In addition, while there is not enough literature on what happened between the 202BC and the 15th Century, a close 
analysis of the literature shows that some African indigenous people had traditional M&E practices, such as the tradition of 
Imihigo in ancient Rwanda back in the 15th Century. These traditional M&E practices included some M&E practices similar to 
what is done today. Furthermore, while ancient M&E practices in the ancient animal and human kingdoms are necessarily 
different from modern (or current) M&E practices, it can be safely concluded that current public policy, programmes, and 
projects M&E practices that have existed since the early 1700 have their roots in M&E practices that existed for many 
millennia ago. For example, many people roast meat on the fire the same way, and the early humans roasted it for the first 
time for the same reasons. The fact that practices similar to programme evaluation were evident in education and public 
health in the 18th Century further proves that M&E practices are not new.

However, the analysis in this paper shows that Ancient M&E practices went through many stages of development, 
refinements and perfection from the Age of Reform/Measurement (1792-1900) to become Results-Based M&E, or Evidence-
Based Evaluation (EBE) in the 20th and 21st centuries, the evidence presented in this paper demonstrates that the genesis of 
public policy, programme and project M&E practices is routed in Ancient African traditions.

Ethical considerations

Not applicable.

Conflict of Interest

The author declares no conflict of interest.

Funding

This research did not receive any financial support.

References

Finnis A (2014) Don’t get out of your depth! Incredible pictures show orangutan using a stick to check river is safe to cross.
Journal of Association of Arab Universities for Tourism and Hospitality 1-16.
Kusek JZ, Rist RC (2008) Ten Steps to a Results Based Monitoring and Evaluation System. Ryan Plummer; Amanda Smits; Samantha Witkowski; Bridget McGlynn; Derek Armitage; Ella-Kari Muhl; and Jodi Johnston.
Murasi II, Auriacome CJ (2022) The role of the imihigo as a performance management instrument in Rwanda.


Accessed on: 13September 2022


Accessed on: November 17, 22.