NEW DEVELOPMENTS

Evaluation of holistic sexuality education: A European expert group consensus agreement

Evert Ketting*, Minou Friele† and Kristien Michielsen‡; On behalf of the European Expert Group on Sexuality Education

*University Medical Centre, Radboud University, Nijmegen, the Netherlands, †Federal Centre for Health Education (BZgA), Cologne, Germany, and ‡International Centre for Reproductive Health, Ghent University, Ghent, Belgium

ABSTRACT

Objectives Holistic sexuality education (HSE) is a new concept in sexuality education (SE). Since it differs from other types of SE in a number of important respects, strategies developed for the evaluation of the latter are not necessarily applicable to HSE. In this paper the authors provide a basis for discussion on how to evaluate HSE.

Methods First, the international literature on evaluation of SE in general was reviewed in terms of its applicability to HSE. Second, the European Expert Group on Sexuality Education extensively discussed the requirements of its evaluation and suggested appropriate indicators and methods for evaluating HSE.

Results The European experience in SE is scarcely represented in the general evaluation literature. The majority of the literature focuses on impact and neglects programme and implementation evaluations. Furthermore, the current literature demonstrates that evaluation criteria predominantly focus on the public health impact, while there is not yet a consensus on sexual well-being criteria and aspects of positive sexuality, which are crucial parts of HSE. Finally, experimental designs are still considered the gold standard, yet several of the conditions for their use are not fulfilled in HSE. Realising that a new evaluation framework for HSE is needed, the European expert group initiated its development and agreed upon a number of indicators that provide a starting point for further discussion.

Conclusions Aside from the health impact, the quality of SE programmes and their implementation also deserve attention and should be evaluated. To be applicable to HSE, the evaluation criteria need to cover more than the typical public health aspects. Since they do not register long-term and multi-component characteristics, evaluation methods such as randomised controlled trials are not sufficiently suitable for HSE. The evaluation design should rely on a number of different information sources from mixed methods that are complemented and triangulated to build a plausible case for the effectiveness of SE in general and HSE in particular.

KEYWORDS: Europe; Evaluation; Indicators; Sexuality education
INTRODUCTION

The past decade has witnessed growing international interest in sexuality education (SE)\(^1\)–\(^5\). A recent advancement has been the development of the concept of holistic sexuality education (HSE) by the European Expert Group on Sexuality Education. HSE is defined as ‘Learning about the cognitive, emotional, social, interactive and physical aspects of sexuality’. Accordingly, ‘Sexuality education starts early in childhood and progresses through adolescence and adulthood. It aims at supporting and protecting sexual development. It gradually equips and empowers children and young people with information, skills and positive values to understand and enjoy their sexuality, have safe and fulfilling relationships and take responsibility for their own and other people’s sexual health and well-being’\(^3\).

This new approach needs to be evaluated in order to test its utility and effect and address the legitimate demand for evidence-based policy.

OBJECTIVES

The main objective of the authors of this article is to initiate the development of a framework for the evaluation of HSE. By this we mean to contribute to improving the quality of evaluation of SE programmes in general and of HSE in particular. In order to achieve these goals, we will:

- Provide general (background) information on the key characteristics of HSE so as to present the subject of a potential evaluation.
- Review dominant practices of SE evaluation and assess their utility for evaluating HSE.
- Suggest alternative and additional evaluation criteria, indicators and research methods for HSE.

Key characteristics of HSE

The concept of HSE refers to an approach to SE that became dominant in recent decades in an increasing number of European countries. HSE has been the result of an evolutionary process that started in Sweden as far back as 1942, when SE was introduced for the first time in schools. Currently, examples of SE programmes that include several core characteristics of HSE are also found in countries such as Sweden\(^6\), Norway\(^7\), the Netherlands\(^8\), Belgium\(^9\) and Estonia\(^10\). The programme titles sometimes explicitly indicate a positive character, like: ‘Long live love’ (Netherlands), ‘Good lovers’ (Belgium), and ‘Love talks’ (Austria).

These programmes are characterised by their fundamentally positive approach towards sexuality, i.e., by considering sexuality as a potential source of joy and happiness and not predominantly as a health risk. Furthermore, they do not aim to prevent young people from starting sexual relationships but accept that young people will engage in loving relationships, which may gradually include sexual elements. HSE thus is conceptualised as guidance throughout life, with the aim of developing and strengthening the ability of learners to make conscious, satisfactory, healthy and respectful choices regarding relationships and sexuality. HSE starts early because what is learnt at a young age has a lasting impact on how sexuality is managed later in life\(^11,12\). It should be emphasised, however, that HSE at younger ages is very different from HSE at older ages. It is crucial to note that programmes should be age and developmentally appropriate\(^14\). For example, young children learn about the proper names of body parts or the meaning of friendship. Later, other themes become relevant, such as the biology of reproduction or ways to prevent pregnancy and sexually transmitted infections (STIs).

HSE-intended outcomes are threefold. They comprise knowledge items (such as knowledge about the human body, its development and functions), personal and behavioural skills (e.g., to be able to build intimate relationships in which there is mutual understanding and respect for each other’s needs and boundaries) and value orientations (e.g., respect for sexual diversity and gender differences).

In a number of ways, HSE differs from its alternatives, i.e., two other approaches to SE which have become dominant in recent decades, particularly in the USA\(^13\): (i) abstinence-only education; and (ii) comprehensive sexuality education (CSE). Abstinence-only focuses on delaying sexual intercourse until marriage. It does not include information and advice on safe sexual behaviour or contraception and it does not usually include issues like emotions, feelings or biological changes during puberty. It has been scientifically demonstrated that abstinence-only does not generate the intended impacts\(^15\).

There is as yet no universally agreed definition of CSE and the concept is interpreted and used quite
differently across the globe, depending on time, country or authors. It varies in practice between ‘abstinence’\textsuperscript{16}, where information on contraception is only added to the abstinence-only approach, to fully comprehensive approaches, such as those developed by the International Planned Parenthood Federation (IPPF)\textsuperscript{1} or by the United Nations Educational, Scientific and Cultural Organization (UNESCO)\textsuperscript{2}. Increasingly, CSE reflects the holistic approach.

The core characteristics of a programme also define the criteria for its evaluation. Current evaluation approaches that were often developed for the purpose of evaluating abstinence-only or the varieties of CSE may be difficult or inappropriate for evaluating HSE because of three core differences: first, HSE is not an intervention but a learning process, spread out over several years; second, HSE does not aim at behaviour change: instead its focus is on behavioural preparation or development; third, HSE starts from general pedagogical and educational theory, and not from theories of behavioural change.

\textbf{METHODS}

First, we performed an assessment of current practices of evaluating SE in Europe in order to study their utility for evaluating HSE. Recent years have seen the publication of two comprehensive review studies on school-based SE that included several European countries\textsuperscript{2,17}. From a total of over 80 studies reviewed, the majority were from the USA, with only six\textsuperscript{17} and nine\textsuperscript{2} originating from Europe. To identify additional studies, we performed a literature search in Web of Science and PubMed, using the following search terms: ‘sex education’ or ‘sexuality education’ and ‘evaluation’ or ‘process’ or ‘implementation’ or ‘effect’ or ‘impact’. We limited the search to publications from 2000 onwards. This search was complemented by studies identified through contacting European experts and by screening reference lists. It was not the intention to present a comprehensive overview or perform a systematic review of available studies; rather, we intended to highlight general trends in current practice and identify strengths as well as shortcomings in their applicability to the evaluation of HSE.

The second part of this article is the result of a consensus-building process within the European expert group, which has been collaborating since 2008 under the auspices of the German Federal Centre for Health Education (BZgA) and the World Health Organization (WHO) Regional Office for Europe. For the purpose of this paper, two meetings on HSE evaluation were convened, complemented by extensive e-mail consultation, after which the paper was accepted as a consensus agreement.

\textbf{RESULTS}

\textbf{Part 1: General trends in current evaluation practice}

The initial literature search generated 1713 titles, which were reduced to 48 studies on some form of evaluation of various approaches to SE in Europe. We identified an additional 72 sources through European experts (often grey literature or in other local languages than English). In the analysis of these papers, we focused on the main types of evaluations, the dominant evaluation criteria, and the prevailing evaluation designs and methods.

\textbf{Evaluation type: Strong focus on outcome evaluations compared with programme and implementation evaluations}

Three types of evaluation were considered: outcome/impact, implementation and programme evaluation. Outcome/impact evaluation assesses the success of the programme in terms of knowledge, attitudes and skills learners have developed as a result of the programme, and ultimately in terms of their behaviour and its possible impact on their health and well-being. Outcomes are the immediate results of the programme; impact is the medium- and long-term effect of it. In implementation evaluation the question is to what extent a programme is developed and implemented as recommended. Programme evaluation aims at studying the programme content, development process and intended mode of delivery. In the current literature, there is a strong focus on short-term outcomes of SE. This is linked to the fact that many interventions or programmes are evaluated immediately after the pilot stage in order to decide whether they should be continued or not. Furthermore, long-term impact evaluations are complex (see Evaluation design, below), as well as very costly, and are therefore seldom found in the literature.
In reporting the results of outcome evaluations, a substantial number of authors refer to problems in implementing interventions or programmes. However, implementation evaluations are rarely published. If published at all, they often solely address specific aspects of the implementation, e.g., compare results of adult-led versus peer-led interventions, or capture the experience of certain facilitators in implementing the programme: teachers, nurses, or, less frequently, the perception of students. We found only a few reports of well-elaborated implementation evaluations that could be applied to the evaluation of HSE.

Moreover, programme evaluation is rarely found in academic literature. If the programme itself is evaluated, it is mainly in the form of a narrow assessment, briefly mentioned in the outcome evaluation report. Sometimes the 17 criteria of successful programmes (such as addressing gender issues, considering community norms, providing quality training to educators) developed by Kirby et al. are used as indications of programme quality. However, these are relatively narrow criteria for programme success and have limited applicability to HSE. In 2013, assessment tools have been developed by UNESCO and IPPF that focus on programme evaluation.

Evaluation criteria: Dominance of public health impact

Evaluation criteria are the standards against which the quality and effects of SE are measured. In terms of outcome and impact indicators, the articles found in the literature search, including the review by Kirby et al. and the UNESCO guidance on SE, clearly indicate a strong focus on public health indicators and behavioural change. Two public health indicators are dominant: pregnancy and STIs/HIV. Dominant intermediate variables considered to influence these biological markers are age at initiation of sex, frequency of intercourse, number of sexual partners, condom use, contraceptive use, and sexual risk-taking in general. Criteria that determine these behaviours are mostly knowledge, skills, attitudes, norms and behavioural intentions.

The significance of these evaluation criteria and indicators is obvious. Questions arise, however, when it comes to their interpretation; for example, whether interventions are considered successful because initiation of sexual activity is delayed, frequency of sex and the number of partners reduced and condom/contraceptive use increased, or because behavioural norms have become more restrictive. Furthermore, although there are exceptions, we noticed very limited use of indicators that focus on positive aspects of sexuality. Even though indicators such as self-efficacy or the ability to communicate about feelings and wishes are used, they are usually not considered important in their own right but with respect to fostering the behaviours just mentioned. Indicators measuring the ability to experience pleasurable and satisfying sexual relationships are hardly ever used.

Programme and implementation quality is mostly evaluated on the basis of the previously mentioned criteria for successful programmes of Kirby et al. or on general national learning objectives. They are mainly treated as intermediate variables influencing the outcome and impact of the programme, and not as critical stand-alone aspects of SE.

Evaluation design: Striving to demonstrate causality

(Clustered) randomised controlled trials (cRCTs), followed by quasi-experimental designs and pre-/post-measurements, still tend to be seen as the gold standard for outcome and impact evaluations, even though their use for evaluating complex public health interventions has been questioned for several years. Indeed, (c)RCTs are well suited to demonstrate a clear cause-and-effect relationship. However, a number of prerequisites are usually ignored when applying RCTs to SE programmes. First, in most SE evaluation studies, study participants are well aware of the group to which they have been assigned (i.e., intervention or control), which may influence the study results and generate reporting bias, and randomisation may be compromised by the fact that, in the end, participants often decide themselves to what extent they participate. Second, in many European countries there are national SE programmes, making it impossible to create an adequate control group of young people who have not been subjected to it. Third, the development of sexual identity is a long and complex process, influenced by many factors, of which SE is only one (others include parents, peers, mass media and sociocultural environment). It is virtually impossible to control for all these factors and identify the specific causal effects of SE in the long term. Finally, good SE programmes start at a young...
age, a moment where there are no relevant knowledge, attitude or behaviour indicators related to sexual intercourse because the children are still very young (i.e., there is no baseline information).

Additional challenges arise in making sure that the evaluation is sufficiently powerful. Especially in populations with low incidences of STIs, adolescent pregnancy or certain sexual behaviours, demonstrating a significant reduction in incidence requires enormous sample sizes, while many trials are underpowered because of a limited estimated sample due to recruitment or financial problems.37

Consequences for HSE evaluation

Taken together, this overview leads to the conclusion that the evaluation methods that have become the international standard are not sufficiently suitable for evaluating programmes that were developed in accordance with HSE principles. Neither do the typical pre-/post-measurements meet HSE’s long-term, multi-component and national characteristics, nor can they assess whether the objectives of HSE have been reached. Since HSE promotes sexual well-being and health throughout the life-cycle, at the moment of ‘baseline’ (ages 6 to 7) there is no relevant knowledge, attitude or behaviour related to sexual intercourse. A new evaluation framework for the evaluation of HSE should meet the following requirements:

- Include a focus on programme content and implementation quality of SE.
- Expand on the exclusive focus on changing particular behaviours, such as the number of sexual partners (in terms of correcting ‘wrong’ into ‘right’ behaviour) and sexual risk reduction, so as to include behavioural preparation and sexual well-being.
- Take into account that HSE is multifaceted and implemented over a period of several years. HSE is not an intervention but an educational process.

Part 2: Recommendations for evaluating HSE – a consensus agreement

Based on the findings and considerations outlined above, the European expert group came to the following recommendations. Evaluation of HSE programmes requires all three types of evaluation: (i) programme; (ii) implementation (or process); and (iii) outcome/impact evaluation, each assessing to what extent the quality criteria of HSE as outlined in the ‘Standards for sexuality education in Europe’3 are met. For each of the three evaluation types, examples of quality or outcome/impact criteria and some related indicators are needed. Quality and impact criteria are defined here as ‘evaluation parameters’ or ‘general characteristics’ that have to be translated into measurable entities called ‘indicators’, in evaluation practice. These criteria and indicators will have to be further developed in the future in order to generate a comprehensive set for evaluation research. These criteria should be based on the characteristics and objectives of HSE, as outlined in the standards3, supplemented by the results of consensus building within the European Expert Group on Sexuality Education. For each of the evaluation types, we also suggest a number of research methods (Figure 1).

Programme evaluation

The criteria for assessing the quality of HSE programmes should be based on the principles of HSE3 and additionally on general recommended characteristics mentioned throughout the standards3. Several of these quality criteria cross through the curriculum. The quality criteria are as follows:

- Positive approach to sexuality as an important human potential.
- Age and developmentally appropriate: the curriculum corresponds to the realities of children’s and young people’s lives at different stages of development, and to their level of interest and understanding as scientifically assessed.
- Culturally and socially responsive: the realities of the existing sexual culture are not taken for granted but are the subject of critical reflection and discussion.
- Gender equality and acceptance of (sexual) diversity: the intention is to prevent or reduce social isolation, intimidation and discrimination based on gender or sexual identity, and to increase self-acceptance.
- Based on a human rights approach, particularly the right to knowledge, to (sexual) self-determination and to equal treatment.
- Based on a holistic understanding of well-being: apart from physical health considerations, psychological, social and interactive aspects of sexuality are dealt with as being relevant in their own right,
and not only as determinants of healthy sexual behaviour.
• Contributing to a fair and compassionate society: empowering individuals and communities; young people are enabled to make their own informed choices and learn to take responsibility for them.
• Providing scientifically accurate information: learners have the right to ask questions and are entitled to appropriate and honest answers.
Furthermore, concerning programme delivery, the main principles include:

- Involvement of children and young people in assessing their own needs, in the development and implementation of lessons, and participation in interactive teaching methods.
- Variety of teaching methods: different objectives require different teaching methods and strategies to be effective.
- Flexible and adaptable to different needs, groups, settings, etc.

Examples of indicators of programme quality include:

- Cross-cutting quality criteria are explicitly mentioned and explained in the programme documentation and reappear regularly in various lessons.
- Positive approach: tone, illustrations and examples in the curriculum are engaging, ‘sexy’, joyful, and not fear-based.
- Age and developmentally appropriate: the curriculum extends over several years and themes are revisited, gradually in more depth, with increasing age.
- Gender equality and acceptance of diversity: lessons which deal explicitly with these subjects are included, appropriate vocabulary is used and all types of sexual identity are included (e.g. by using the term ‘partner’ instead of man/woman).
- Involvement of young people: mechanisms are in place that ensure meaningful participation on all levels of programme development and implementation.

Evaluation methods. The quality criteria can be studied through document analysis and curriculum review.

Implementation evaluation

Implementation (or process) evaluation looks at discrepancies between how a programme is supposed to be implemented in theory and how it is done in practice. This type of evaluation is rather rare in the literature17. There are four areas that need to be addressed in implementation evaluation: (i) process of programme development; (ii) teacher/educator training and support; (iii) linkages with relevant sexual health and well-being services; and (iv) curriculum delivery.

Process of programme development. Even a completely standardised HSE programme needs adaptation to local conditions. How to do this has been the subject of a follow-up publication to the standards3, ‘Guidance for implementation’4. Implementation quality criteria for programme development are that the following activities have been carried out:

- Mapping of challenges and opportunities for introducing HSE.
- Definition of learning objectives for successive age groups.
- Development of lesson plans with defined topics, learning outcomes, methods and suggested means of students’ evaluation/feedback.
- Interdisciplinary development of the HSE syllabus and lesson plans per age group.
- Development of teacher manual and student materials.
- Involvement of young people, parents, community representatives and educational specialists in the entire process.
- Special attention to vulnerable groups with higher levels of risky behaviour and special needs.

Indicators for programme development follow logically from these quality criteria; the successive steps can easily be translated into a checklist of necessary activities.

Teacher/educator training and support. Training of school teachers and other educators is essential for successful implementation of HSE2,16. However, such training is rarely included in pre-service teacher training curricula and many therefore depend on postgraduate training. There are, however, exceptions. In Estonia, SE teacher training was integrated into the human studies training curriculum of two universities in 1996 and 200235, and, more recently, in the Netherlands38 and Finland. Aside from training, teachers also need the support of their school head and school administration, because SE can be a subject that is frowned upon by other teachers or by parents of learners. Finally, teachers need access to background information and suggestions or instructions on how to deliver the different lessons.

Indicators for evaluation research on teacher training and support include:

- Teachers and other educators have been trained on how to deliver the curriculum.
• The school head and school administration take a shared responsibility for HSE.
• A manual on the HSE curriculum is available for educators.

Linkages with relevant sexual health and well-being services.
Linkages between the HSE programme and health, welfare and other relevant organisations in the community are needed for at least two reasons. First, these organisations are important as supportive resources for lectures on specialised issues and subjects that teachers consider to be too sensitive or challenging to teach themselves. Second, learners need to know where they can access the different health and social services that they require, including sexual and reproductive health (SRH) services. In several European countries, the curriculum includes a class visit to a youth health centre, which lowers the threshold to contact it in the future.

Indicators for evaluation research on linkages with services may include:
• The presence of formal collaboration arrangements between the HSE programme and relevant information and service delivery services in the community.
• The involvement of relevant services/institutions in the implementation of the programme, in a variety of ways.

Curriculum delivery. The last and probably most important implementation quality criteria relate to discrepancies between the HSE programme on paper and the way it is actually implemented. A common discrepancy is that the curriculum is only partly implemented. This can be due to lack of time, prioritisation of other subjects, or other reasons. Often the most sensitive themes are omitted because the teacher feels unable to handle them. It is also possible that HSE is optional and is not chosen by all learners. Finally, there may be a wide range of discrepancies between intended and actually implemented lessons.

Indicators for curriculum delivery quality may be:
• The full curriculum has been implemented as intended.
• The curriculum is mandatory for all learners.
• A checklist is used to assess discrepancies between how the curriculum is supposed to be and how it is actually implemented.

• The results of this assessment are used for future programme strengthening.

Evaluation methods. Implementation evaluation requires intensive and time-consuming research efforts. For assessing implementation, relevant evaluation methods that will provide valuable insights are document analysis of the programme development process complemented by interviews with key stakeholders involved in the process. Other relevant evaluation tools are: participant observation, interviews with educators, assessment of linkages with services, and a checklist to identify discrepancies between the curriculum and what has actually been taught.

Outcome and impact evaluation
This type of evaluation looks at what the HSE programme has brought about among learners in the short- (outcome) and longer-term (impact). HSE programmes do not focus on changing but on enabling young people, and therefore behavioural changes cannot be impact criteria. The function of the HSE programme is not to prescribe or impose ways of thinking and behaviour but to facilitate the development of critical thinking and the ability to make conscious behavioural choices. The ultimately intended impact of HSE is that it has facilitated the development of respectful and satisfactory relationships and sexual life. This implies that there are differences in emphasis compared with current mainstream evaluation practice, particularly from those with an exclusive public health approach.

The following three conceptual pillars are critical to impact evaluation of HSE:
• Children and young people themselves are in the best position to evaluate the programme in terms of ultimate impact.
• The meaning of this impact can differ, depending on age, gender, sexual orientation, and other characteristics of young people.
• The impact is broader than physical health, as it also includes psychological, social, cultural and interactive aspects of health and well-being.

Relevant short-term outcomes are those that are known to contribute to this ultimate impact, based on scientific evidence. In most SE programmes, with a focus on safe sexual intercourse and short-term ‘project’ implementation, the relevant impact can only be
evaluated some years after the curriculum ended. That means that for most learners there will be a time lag before they actually start having sexual relationships. Because HSE starts at a young age and runs for several years, outcomes should be evaluated during implementation and not only at the end of it. Outcomes can be measured after each developmental stage of 3 years, as described in the standards of evaluation by children and young people. An often neglected crucial outcome is the appreciation of the curriculum, and the way it is delivered, by the learners themselves. This aspect is important because if learners perceive (H)SE as meaningful, interesting and critically stimulating it is more likely to have a lasting impact on them. Examples of quality criteria of the curriculum itself and its modes of delivery that can easily be transformed into indicators for questionnaires to be used with learners are:

- It is felt to correspond to the realities of learners’ lives.
- It responds to their curiosities and questions.
- The materials are relevant and visually interesting.
- Learners feel motivated and interested to participate in these classes.
- It provides new scientifically accurate or more detailed information than they already have.
- It empowers and enables learners to take responsible decisions.
- Learners are positive about the way the teacher handles the lessons.

Outcomes. The standards describe a set of key learning outcomes of HSE:

- Knowledge about the human body, its functions and processes, in particular regarding sexuality and its various dimensions, and about protection against health hazards of sexual behaviour.
- Practical knowledge about available and reliable sources of information, advice, counselling and medical services related to sexuality.
- Tolerance and respect towards various (sexual) lifestyles, gender identities and sexual orientations, attitudes and values.
- Critical reflection on norms and values relating to sexuality and human rights.
- Ability to develop as a sexual being, to express feelings and needs, to communicate and negotiate intimate relationships and sexuality, to experience sexuality in a pleasurable manner and to develop one’s own gender and sexual identity.
- Empowerment for making informed choices, observing sexual rights and being in control of one’s sexual health.
- Responsible behaviour towards oneself and one’s partner.
- Positive attitude to gender equality.

These outcomes can be defined and measured at different ages. In evaluation practice, this short overview of core areas of outcomes needs to be elaborated in more detail. Unfortunately, providing a complete list of indicators is beyond the scope of this paper and only some examples can be suggested here. There are three types of outcomes: those related to knowledge, to attitudes and values, and to (behavioural) skills.

Indicators for knowledge can immediately be derived from the curriculum contents and are easy to score as knowledge items (e.g., being able to name all body parts correctly) in examinations. Indicators for attitudes and values are more complex. For example, outcomes such as the development of tolerance and respect require the use of scales. For some concepts these scales exist, for others they have to be developed and validated. The same applies to indicators for skills development. For example, ‘the ability to communicate about sexuality issues’ cannot be measured directly. It is known that it requires positive self-identity, self-esteem and self-efficacy, for which measurement scales have been developed and validated. Such scales can be used as indicators, if necessary after adaptation to the age of the respondents.

Impact. The overall impact of the programme should logically be determined by the programme outcomes. For example, a skill (outcome) has been developed, and this will contribute to a corresponding behaviour (impact).

As mentioned earlier, most impact variables used in current mainstream SE evaluation, such as those related to sexual intercourse, often only become relevant several years after the end of the programme.

Impact variables that are directly related to having intercourse are irrelevant when adolescents are not yet sexually active. In Europe, the vast majority of young people start having intercourse after the age of 15; only about a quarter of them are already sexually active by the age of 15. This means that the first sexual
intercourse usually happens several years after they have been educated on safe sexual behaviour. This explains why most evaluation research uses ‘intended behaviour’ instead of ‘actual behaviour’ as a success indicator, because the latter is in practice much more difficult to assess. Because there are also other important factors that complicate impact assessment (see earlier), outcome evaluation, which can be done immediately, should be given more attention than impact evaluation. It furthermore makes sense also to develop impact variables that are relevant at other stages of development of children and adolescents.

As far as impact evaluation is realistically possible, the core long-term HSE impact variable is ‘young people are prepared for consensual, voluntary, equal, pleasurable, safe and satisfactory relationships and sexual lives’. Because this is a positive impact, indicators have to be developed and used that measure sexual life in positive, instead of negative, terms. Fortunately, there is a new tendency in research to look at sexuality from a positive perspective and to evaluate the impact of positively experienced sexuality on general well-being.46

Indicators that reflect the intended impact of HSE include:

- Relationships and sexual contacts result from mutual consent and individuals express self-confidence in their ability to make choices in relationships.
- Relationships and sexual contacts are experienced positively, pleasurably.
- Individuals demonstrate satisfaction with their ability to make and act on SRH choices.
- Individuals demonstrate tolerance and respect in relation to gender equality and sexual orientation.

At the same time, it is essential that impact evaluation in terms of avoidance of health hazards remains on the agenda. This includes three well-known and extensively evaluated indicators:

- Unintended pregnancy.
- STIs/HIV infection.
- Sexual coercion, abuse and violence.

Evaluation methods. For the assessment of outcome/impact, suggested evaluation methods include:

- Qualitative methods, such as interviews and focus group discussions with learners, educators and parents (especially for the younger children) on their perceptions and experiences with receiving/giving HSE.
- Case-study approaches to assess barriers and facilitating factors.
- Population-based surveys on the evolution of, for example, attitudes, satisfactory sexual lives and biological markers (such as pregnancy or STIs) in the reference populations10,50.
- Quasi-experimental designs for well-aligned outcomes such as knowledge.
- Cross-sectional surveys among key stakeholders on the importance of SE.

Complementarity and triangulation in evaluation types and methods

Given the many purposes of HSE, its multidimensionality and its long-term, context-specific and often national character, using a probability-based design to demonstrate causal effects between HSE and certain outcomes/impacts is virtually impossible. At the same time there is no clear-cut alternative. In alignment to the ongoing discussion on the use of (c)RCTs in evaluating complex programmes, we agree with Laga et al.32 when they advocate ‘realism and pragmatism when it comes to generating more convincing evidence to guide prevention programming’. Such realist evaluation recognises that sexuality is influenced by multiple interlinked factors, acknowledges that programmes work in different ways for different people, and takes into account the context in which a programme is implemented47,48. Where probability evaluations ask ‘Does the programme work?’, realist evaluation studies ask ‘What works for whom, in what contexts, and how?’.

Instead of focusing on a single evaluation design, we propose a blended approach that relies on a number of different information sources from mixed methods. These methods are to be complemented and triangulated to build a plausible case for the effectiveness, utility and necessity of SE49. Such triangulation is not only necessary between different evaluation methods but also between evaluation types: an outcome or impact evaluation should be interpreted in the light of a programme and implementation evaluation.

CONCLUSION

HSE primarily aims at preparing people for consensual, voluntary, equal, pleasurable, safe and satisfactory
relationships and sexual lives. In order to learn from experience and improve HSE programmes, it is crucial that HSE is evaluated. However, given the strong and often exclusive focus on public health in current evaluation research, indicated by our literature search, a new evaluation framework for HSE is required.

With this article the European Expert Group on Sexuality Education is initiating the development of such a framework. As outlined above, we propose a number of quality criteria and indicators that align with the specificities of HSE and take on a realist approach that acknowledges that programme development, content and delivery are of equal value to envisaged outcomes. Recognising other strong influences on children’s and young people’s sexuality, and taking into account the characteristics of HSE, we argue that it is not possible to demonstrate causal effects of HSE using only one evaluation technique. Rather, we suggest complementing and triangulating different data sources in order to build a plausible case for its merits, looking at the medium- to long-term outcomes in relationships, sexual well-being and good health. The logical next step will be the development of a full set of indicators for programme development and implementation of HSE, as well as for its intended outcomes and impact.

ACKNOWLEDGEMENTS

This paper was jointly developed by the above-mentioned authors and the following additional members of the European Expert Group on Sexuality Education (in alphabetical order):

- Dan Apter (Sexual Health Clinic Väestöliitto, The Family Federation of Finland, Finland)
- Vivian Barnekow (WHO Regional Office for Europe, Denmark)
- Doortje Braeken (IPPF, UK)
- Raisa Cacciatore (Sexual Health Clinic Väestöliitto, The Family Federation of Finland, Finland)
- Marina Costa (Youth and Sexuality Education Service ‘LustundFrust’, Switzerland)
- Karolien Dekker (IPPF European Network, Belgium)
- Sanderijn van der Doef (Rutgers WPF, the Netherlands)
- Lyubov Erofeeva (Russian Association for Population and Development, Russia)
- Erika Frans (Sensoa, Belgium)
- Joanna Herat (UNESCO, Division of Education for Peace and Sustainable Development, Section of HIV and Health Education, France)
- Angelika Hessling (BZgA, Germany)
- Olaf Kapella (Austrian Institute for Family Studies, University of Vienna, Austria)
- Merike Kull (Department of Sport Pedagogy, University of Tartu, Estonia)
- Margareta Larsson (Department of Women’s and Children’s Health, University of Uppsala, Sweden)
- Gunta Lazdane (WHO Regional Office for Europe, Denmark)
- Olga Loeber (European Society for Contraception and Reproductive Health Care, the Netherlands)
- Ulla Ollendorf (Norwegian Directorate of Health, Ministry of Health, Norway)
- Kai Part (Department of Obstetrics and Gynaecology, University of Tartu, Estonia)
- Simone Reuter (Contraception and Sexual Health Service, Sherwood Forest Hospitals NHS Foundation Trust, UK)
- Marija Vasileva-Blazev (United Nations Population Fund, Regional Office for Eastern Europe and Central Asia, Turkey)
- Ineke van der Vlugt (Rutgers WPF, the Netherlands)
- Christine Winkelmann (BZgA, Germany)

We would furthermore like to thank Birgit Dederichs-Bain for facilitating the expert group’s consecutive workshops, and the members of the expert group not mentioned above.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES


50. Apter D. Recent development and consequences of sexuality education in Finland. BZgA Forum Sex Educat Fam Plann 2011;2:3–8.