frontiers in Ecology and Evolution

METHODS
published: 22 December 2017
doi: 10.3389/fevo.2017.00165

More Pitfalls with Sperm Viability Staining and a Viability-Based Stress Test to Characterize Sperm Quality

Barbara A. Eckel, Ruijian Guo and Klaus Reinhardt*

Applied Zoology, Department of Biology, Technische Universität Dresden, Dresden, Germany

Sperm viability (SV), the proportion of live sperm in a sample, is a widely applied measure of sperm quality but few studies test its robustness. At least three reasons make SV problematic as a surrogate for sperm quality. First, reviewing the ecological literature revealed that previously identified methodological pitfalls have not been overcome, including low cross-study standardization of protocols, inadequate statistical treatment, and unaccounted for within-sample heterogeneity. Second, SV is affected by biological variation such as between species, reproductive organs, or sperm age cohorts. Third, the proportion of live sperm extracted from males appears more related to male than to sperm quality in the sense of the future performance of sperm. We propose an alternative method to assess sperm quality by characterizing the temporal decrease of SV in a stressor medium and illustrate in two species, the common bedbug (*Cimex lectularius*) and the fruit fly (*Drosophila melanogaster*) how some common methodological pitfalls may be circumvented. Our data empirically support the well-known but little-considered facts that (i) non-blind measurements may alter SV and (ii) that SV frequently have non-significant repeatability within one sample. (iii) Cross-sectional sampling of ejaculates showed that this heterogeneity even masked a biological pattern—the sperm stratification within males. We show (iv) that this shortcoming can be overcome by following the temporal decline of SV of a sperm subsample in a stress test. Finally, (v) comparing the staining pattern of sperm between *Cimex* and *Drosophila*, we found that in the latter, the visibility of sperm is substantially delayed (30 min) when sperm density is high. We show that this delay in stained sperm visibility was, however, not biased toward dead or live sperm. To measure sperm quality, we advocate analyzing the temporal decline in SV in a stressor medium over current protocols that use SV *per se* and blinding samples for SV measurements. As cell viability is widely used in biological and medical laboratory studies, our protocol may be useful to characterize cell quality beyond ecology and evolution.

Keywords: live/dead kit, membrane integrity, propidium iodide, sperm senescence, sperm stratification, sperm viability, SYBR14

OPEN ACCESS

Edited by:
Jordi Figuerola, Estación Biológica de Doñana (CSIC), Spain

Reviewed by:
Anders Pape Moller, Centre National de la Recherche Scientifique (CNRS), France
Francisco Garcia-Gonzalez, Estación Biológica de Doñana (CSIC), Spain
Fabrice Helfenstein, University of Neuchâtel, Switzerland

****Correspondence:***
Barbara A. Eckel
barbara.eckel@tu-dresden.de

Specialty section:
This article was submitted to Behavioral and Evolutionary Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *10 August 2017*
Accepted: *08 December 2017*
Published: *22 December 2017*

Citation:
Eckel BA, Guo R and Reinhardt K (2017) More Pitfalls with Sperm Viability Staining and a Viability-Based Stress Test to Characterize Sperm Quality. Front. Ecol. Evol. 5:165. doi: 10.3389/fevo.2017.00165

INTRODUCTION

Sperm quality is routinely measured in evolution and ecology (Holman, 2009), but also in reproductive medicine (World Health Organization, 2010), reproductive toxicology, animal breeding, and aquaculture (Morrell and Rodriguez-Martinez, 2009). In ecological and evolutionary studies, sperm viability (SV) is often used as a proxy for sperm quality that influences sperm competition (Parker, 1970; Parker and Pizzari, 2010) beyond a mere difference in sperm number. The number, or sometimes the proportion *per se*, of live sperm in an ejaculate is considered evolutionarily significant (Garcia-González and Simmons, 2005). Across insect species, polyandrous species contained a higher proportion of viable sperm than related monandrous species and it was suggested that SV evolves with sperm competition (Hunter and Birkhead, 2002). SV has experienced a surge in usage as commercial kits became available that measure SV in the form of the stability of the sperm membrane. Briefly, the DNA in sperm heads of intact cells is stained with a green fluorescent dye (SYBR14®) that is replaced by a red fluorescent dye (propidium iodide, PI) as the membrane becomes leaky.

However, measuring sperm quality by current protocols of SV staining is not without pitfalls. For example, defining SV by the number of live sperm per ejaculate is a mere refinement of sperm number and better seen as a male parameter, rather than a parameter of future sperm performance. For SV to represent a sperm quality parameter beyond the live or dead dichotomy, it would be necessary to show that a correlation exists between the numbers of dead sperm in an ejaculate and the prospective mortality rate of the living sperm in that ejaculate (Reinhardt, 2007). This relationship does not seem to have been demonstrated. In the following contribution, we will therefore, propose a measure of sperm quality that has some prospective component, the temporal decrease in SV in a stressor medium. In addition to these conceptual issues, Holman (2009), reviewing the use of SV in ecology and evolution, identified additional technical problems: (1) because the SV assay itself causes sperm mortality, the "true" number of live and dead sperm cannot be known. (2) Given the wide-ranging environmental effects on sperm quality (Reinhardt et al., 2015), sperm from different organs or at different cellular ages may have different membrane properties. Comparing their SV may therefore be confounded by different amounts of sperm being killed during dissection and during staining. Given the rapid evolutionary change in sperm form (Pitnick et al., 2009) and sperm function (Reinhardt et al., 2015), different species may differ in both in SV as well as its susceptibility to the staining dye. (3) Both in nature and on the microscopic slide, SV may not be independent of the total number of sperm.

To attenuate these and other problems, several procedural recommendations were made (Holman, 2009): (i) to block treatment groups, (ii) to record data in a blind way, (iii) to measure the repeatability of SV in a sample, (iv) to count an unbiased proportion of sperm, (v) to simultaneously measure sperm number and viability and, (vi) to analyse viability with binomial generalized linear models with a logit link function.

Adding to these specific concerns of SV measurement, a string of papers in the ecological literature re-iterated old (Milinski, 1997) and recent concerns of observer bias (Holman et al., 2015; Forstmeier et al., 2016). The most important way to circumvent observer bias is blinding the sample IDs to the observer (Holman et al., 2015; Forstmeier et al., 2016). Unless this is done, unconscious psychological biases can lead to false-positive results (Forstmeier et al., 2016). For example, effect sizes in matched pairs of non-blind and blind ecological and evolutionary studies were substantially higher in the non-blind studies (Holman et al., 2015). Moreover, thus reported exaggerated effect sizes exceeded the effect sizes typically found in evolution and ecology (Holman et al., 2015). Because we found that studies citing Holman (2009) and using SV staining were not exempt from these concerns (see Results), we assessed some of the concerns empirically. We present a method to assess sperm quality that is based on the temporal decrease in SV of sperm placed in an osmotic stressor. The method circumvents effects of within-ejaculate heterogeneity which we illustrate by applying it to the task of detecting relatively small differences in SV: We will ask whether within a male, sperm are stratified by cell age from the testes toward the ejaculation site, or whether sperm of all age cohorts are mixed in the male sperm store (Reinhardt, 2007). If sperm are stratified by age, the oldest spermatozoa (with the most strongly damaged membrane and hence, low SV) are predicted to be closest to the ejaculation site (=caudal part of the male sperm store). Younger spermatozoa with fewer exposure to potential damage and with intact membranes will be closest to the production site (=cranial part of the male sperm store) and are predicted to have high SV. Specifically accounting for the biologically unknown "true" value of SV, we here test the prediction of the stratification model that sperm extracted from the cranial part of the male sperm store (the seminal vesicle) have higher quality in the form of a slower sperm aging in a stressor medium than sperm extracted from the caudal site. The sperm mixing model serves as the null hypothesis that there is no effect of sperm collection site on SV and sperm quality.

To test these predictions and to address some of the above-mentioned pitfalls, we measure SV in the common bedbug, *Cimex lectularius,* and re-assess sperm quality in a refined protocol in fruit flies, *Drosophila melanogaster*.

METHODS

Methods of Sperm Viability Staining in Ecology and Evolution

We first established how the methodological suggestions by Holman (2009), had been addressed in ecology and evolution. Until April 2017, 38 studies had cited Holman's article (Holman, 2009) on google scholar (all were also listed in the Web of Science). From these, we collected information on the study species, the compartment from which sperm was extracted, protocol details such as buffer and staining duration, and statistical details. We excluded reviews and restricted our analysis to those 26 studies that used the commercially available SV staining based on SYBR14® and PI.

Empirical Study

Insect Maintenance

We maintained bedbugs (*C. lectularius*) as previously described (Reinhardt et al., 2003). All males used in the experiments were virgins kept in single 15-ml transparent plastic tubes equipped with a piece of filter paper. Males were kept singly to prevent them from removing old sperm by male-male matings (Ryne, 2009). Wildtype *D. melanogaster* (Oregon-R strain) were maintained on standard yeast food at 25°C and 65% RH on a 12:12 h L:D cycle. All males used in the experiments were virgins.

SV Assessment

We stained sperm with SYBR14® (1:50 in DMSO) and PI (LIVE/DEAD® Sperm Viability Kit, ThermoFisher Scientific). Pilot trials were used to reveal the minimum concentration that still provided a clear fluorescence signal under our microscope. We measured sperm viability directly after staining without any further incubation, or at pre-assigned time intervals in the time series protocol (see below). The time between dissection and the first image taken never exceeded 1 min. We took pictures of the stained sperm with a fluorescence microscope (Leica DM5000 B; Leica DMi8, Leica live/dead filter set, Leica DFC 450 camera). We counted sperm on JPEG files by eye and recorded their number using the automated counting function in ImageJ (Schneider et al., 2012). Green sperm were considered alive, red and red-green double stained sperm as dead.

Non-blind SV Measurement in Bedbugs

We dissected the seminal vesicles of 14 virgin males 5–30 days after isolation from their colony in 30 μl phosphate buffered saline (1xPBS) (**Figure 1**). As a sperm stress test, one seminal vesicle was transferred to double-distilled (dd) H_2O, split in the middle using a scalpel, and the cranial and caudal parts of the vesicle were placed into separate drops of 50 μl ddH_2O. Sperm and water were mixed by pipetting them up and down five times. The other seminal vesicle was handled in the same way. The cranial and caudal parts of the first seminal vesicle were stained immediately (time zero, t_0) with 2 μl SYBR14® and 1 μl PI and a cover slip added. We took four pictures of haphazardly chosen areas at 100 x magnification straight away. The sperm of the second vesicle remained in the ddH_2O for 15 min before it was analyzed (t_{15}).

One person who was aware of the sperm stratification hypothesis (but not of the later blind repetition of the study), did the dissections, stained the sperm, selected the count areas in the microscope and counted the sperm. The results were strongly supporting the hypothesis (see Results) and we decided to repeat the experiment in a blind test.

Blind SV Measurement in Bedbugs

We dissected 15 male bedbugs 76 days after isolation from the colony in 50 μl ice-cold Grace's insect medium (**Figure 1**). We chose this deviation from the original protocol in male age in order to increase the age difference between the sperm cohorts during sperm stratification, and therefore, to increase the experimental effect. We partitioned the vesicle in the middle and squeezed sperm carefully out of either the cranial or the caudal half of the vesicle using forceps and applied the sperm stress test. We transferred 10 μl of sperm immediately into a 1.5 ml Eppendorf tube containing 50 μl sterile ddH_2O. We left the other half of the seminal vesicle in Grace's insect medium for five more minutes and then handled it the same way as the first part. We mixed each sample by pipetting it up and down five times, transferred 10 μl of it onto a fresh microscopic slide and immediately stained it with 0.5 μl SYBR14® and 1 μl PI. After adding a 22 × 22 mm coverslip, four pictures of haphazardly chosen areas were taken straight away of each sample at 200x magnification (t_0). We stained another subsample of 10 μl for 30 min after mixing sperm and ddH_2O (t_{30}) as described above and again took four haphazard pictures. We chose to measure SV after 30 min instead of 15 min as we did not find the decrease in SV in the 15-min period (see Non-blind SV measurement in bedbugs above; see Results).

In half the males, the cranial part of the vesicle was tested first, in the other half the caudal part. The experiment was done blind. B.A.E. did the dissections and staining of sperm while a co-worker unaware of the treatment selected the count areas in the sperm sample to take pictures from. Because often there was no temporal decline or even an increase in SV, it appeared as if the heterogeneity in the sperm sample was large and required a protocol modification. This is specified below.

Blind SV Measurement in Bedbugs with a Time Series Protocol

We dissected 20 virgin male bedbugs 33 days after isolation from the colony as described above, using either the cranial (10 males) or the caudal half (10 males) of the seminal vesicle (**Figure 1**). Directly following the sperm staining (see above) one picture was taken in a haphazardly chosen area (t_0) and the fluorescence excitation source switched off immediately thereafter. The slide was left on the microscope and a picture of exactly the same area was then successively taken at 5, 15, and 30 min after the first picture. During these trials, the excitation source was switched on for ∼30 s per picture. Picture areas were chosen blind with respect to the location the sperm was taken from, by a person unaware of the research question. This protocol also allowed us to score retrospectively whether the number of visible stained sperm (dead or alive) stayed constant over time (corresponding to variation incubation time).

Blind SV Measurement in *Drosophila*

We dissected fifteen 14-day old virgin males in 10 μl of ice-cold Grace's medium. We transferred one seminal vesicle to another 10 μl of Grace's medium and released the sperm from the organ by puncturing it with an insect pin. Almost all sperm, along with 2 μl Grace's medium were transferred into a drop of 10 μl ddH_2O. All was mixed by gently pipetting it up and down six to seven times. Sperm was stained immediately (t_0) with 0.5 μl SYBR14® and 1 μl PI, a 22 × 22 cover slip added and four pictures of haphazardly selected areas were taken at 400x magnification. The second seminal vesicle was treated in the same way except that it was kept in ddH_2O for 30 min before assessing

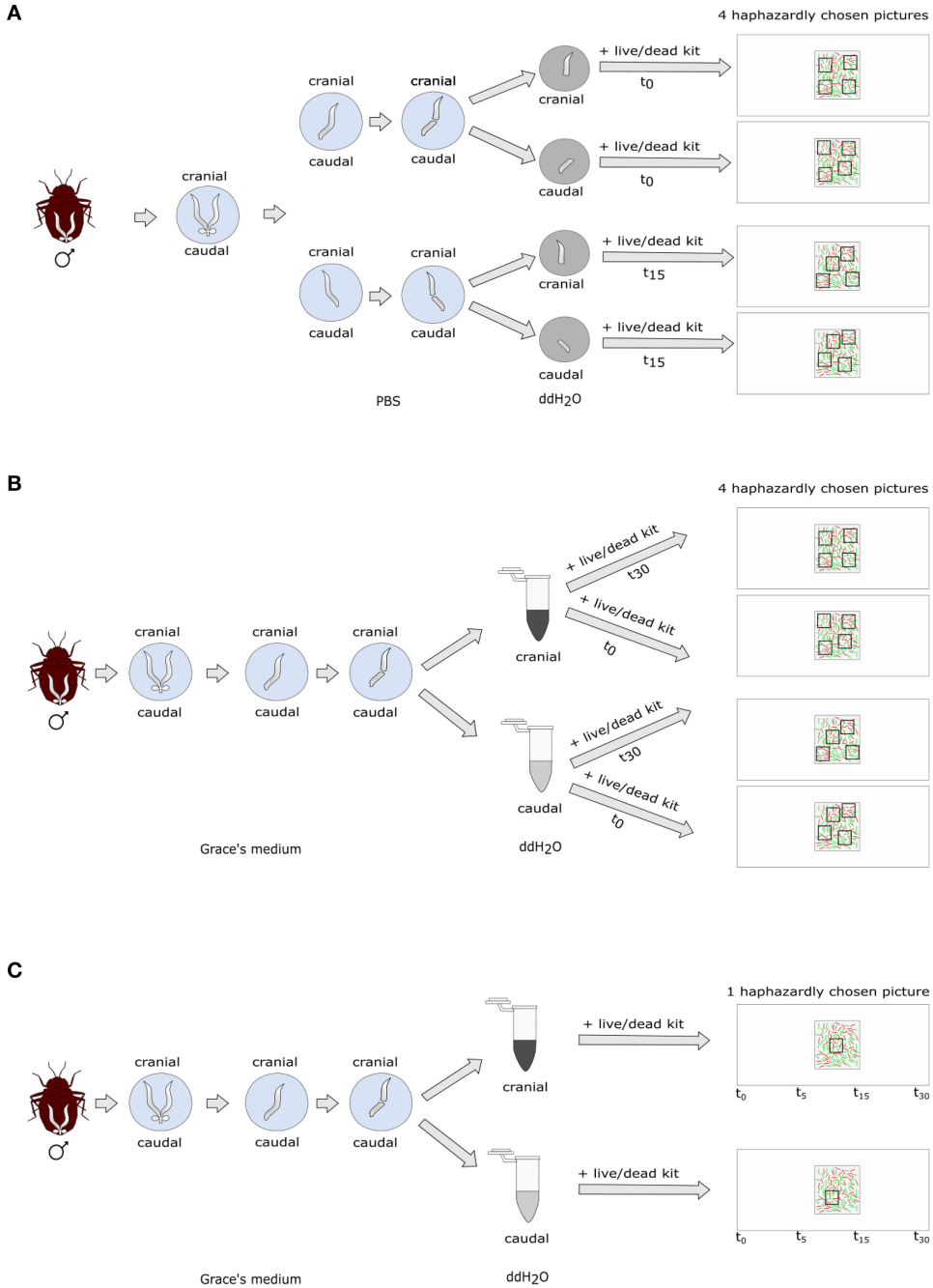

FIGURE 1 | Experimental protocol of **(A)** non-blind standard sperm count; **(B)** blind standard sperm count; **(C)** blind time series count.

SV (t_{30}). The pictures were taken by a co-worker unaware of the t_0 or t_{30} treatment.

Blind SV Measurement in *Drosophila* with a Time Series Protocol

Ten 14-day old virgin males were dissected in 10 µl Grace's medium. Sperm from one seminal vesicle was stained as described in the previous paragraph. One area was haphazardly selected and one picture taken immediately (t_0). The fluorescence excitation source was switched off, and successive pictures of exactly the same area were taken 5, 15, and 30 min after the first picture. The excitation source was switched on for ∼30 s per picture. The area was selected and the pictures taken blind with respect to treatment, by a person unaware of the research question. This protocol also allowed us to score retrospectively whether the number of visible stained sperm (dead or alive) stayed constant over time (corresponding to variation incubation time).

Statistical Analysis

All analyses were performed in R, version 3.3.2 (R Development Core Team, 2016). We analyzed a weighted SV with the *cbind* function (number alive| number dead) using generalized linear mixed models (GLMM; binomial error structure with a logit link) using the lme4 package (Bates et al., 2014), correcting the model for overdispersion (Browne et al., 2005) and pseudoreplication. Full models containing time, location and their interaction (*Cimex*) or time only (*Drosophila*) were reduced in a stepwise backwards mode using the *anova* function to select the final model. SV in the cranial and caudal parts of the seminal vesicle of bedbugs were compared using Welch's two sample *t*-tests for each time point in the time series experiment.

The repeatability of SV across the four haphazardly chosen pictures was analyzed with the ICC package (Wolak and Wolak, 2015) using the intra-class correlation coefficients (ICC) of 10 randomly selected pictures of both fruit fly and bedbug sperm samples, each counted four times. We also analyzed the precision of the counting procedure itself using ICC by counting the sperm on 10 images four times each.

RESULTS

Methods of Sperm Viability Staining in Ecology and Evolution

Eleven out of 26 studies (42%) citing Holman (2009) (**Table 1**) used either blind measures, flow cytometry, or counted the entire sperm sample. Five studies (19%) used the recommended binomial generalized linear or mixed models with a logit link function and six (23%) took repeated measurements of males. Two of the latter found high repeatability of SV, four did not report it. Four studies (15%) used a predetermined number of pictures (five) per sample to evaluate SV but none of them examined the repeatability of SV across these pictures. Accordingly, the heterogeneity within males or within the sample seems unknown for most species. Studies differed widely in the buffers used, in the concentrations and incubation time of SYBR14® and PI employed, even in the same species (**Table 1**). In summary, a minority of studies citing reference (Holman, 2009) would consider some of its recommendations, no study considered all.

Sperm Survival Examined with Cross-Sectional vs. Longitudinal Sampling

We measured empirically the impact of some of the methodological effects identified for SV measurements in the literature (Holman, 2009) in two species, the bedbug *C. lectularius* and the fruitfly *D. melanogaster*. In our samples, we found no correlation between the total number of sperm counted and SV in any of 10 cases (**Table 2**).

Sperm Viability in the Bedbug

Non-blind vs. Blind Measurement

Strongly supporting the sperm stratification hypothesis, SV was significantly higher in the cranial compared to the caudal part of the seminal vesicle (**Table 3**, **Figure 2**) in the non-blind measurement. In the blind measurement, however, there was no evidence whatsoever that SV differed between the cranial and caudal part of the seminal vesicle (**Figure 2**). In both scenarios, SV did not show a significant decrease with time in ddH_2O in the cranial or the caudal part. The surprising lack of sperm mortality may, in principle, be caused by large heterogeneity in SV within a sample. Confirming our suspicion, the intra-class correlation coefficients of SV across four pictures of the same subsample was indeed low to absent at both points in time (**Table 4**). This result was not caused by low repeatability of the counting procedure itself, which was very highly repeatable (ICC = 0.96; 2.5% CI = 0.91, 97.5% CI = 0.99).

Time Series Blind Measurement

The total number of sperm stayed the same over the protocol duration, indicating immediate and complete visibility of bedbug sperm after SV staining (**Figure 4**). SV was found to decline over time in both the cranial and the caudal compartment (**Table 3**, **Figure 3**). The significant location x time interaction effect on SV showed that SV decreased faster in sperm from the cranial than the caudal part.

Sperm Viability in the Fruitfly

Blind Standard

The ICC of SV across pictures was as low as in *Cimex* (**Table 4**), again despite very high repeatability of the procedure itself (ICC = 0.99; 2.5% CI = 0.98, 97.5% CI = 0.997) (pictures taken immediately after staining). Overall, the variation in cross sampled pictures within a sample did not mask the decline of SV over time (**Table 5**, **Figure 3**).

Time Series Blind Measurement

The time series measurement revealed two differences in SV staining between *Drosophila* and *Cimex*. First, the total number of visible sperm increased over protocol duration (**Figure 4**) in *Drosophila* (but not in *Cimex*) indicating incomplete sperm visibility over protocol duration. Second, the delayed visibility was only observed in samples with high sperm density, samples with low sperm density showed constant sperm number over time (Supplementary Figure 1). Both observations prevent protocol standardization across the two species. However, simultaneously they allowed us to carry out an analysis that would not be possible otherwise, namely assessing the influence of delayed visibility on the measurement of SV. We split our dataset into males whose sperm number stayed constant and those whose sperm number increased with protocol duration. Both groups of samples showed virtually identical declines in SV (Supplementary Table 1), indicating that sperm visibility was not biased toward either live or dead sperm (**Figure 5**).

DISCUSSION

Our main aim was to present a sperm stress test as a method to characterize the sperm quality in the form of future sperm performance. We developed a longitudinal approach that had the advantage to account for within-sample heterogeneity and further allows the assessment of a change in visibility of stained sperm. In our case, the sperm visibility of sperm was

TABLE 1 | Parameters of sperm viability measured in studies citing (Holman, 2009) till April 2017.

Species	Organ that sperm was removed from	Amount of sperm suspension mixed with kit	Sperm diluent	Experimental context	Amount/ concentration of SYBR14	Amount/ concentration of PI	Incubation time	Methodological details	Treatment of red and green double- stained cells	Statistical analysis	References
Acheta domesticus (House cricket)	Spermatophore	5 µl	Beadle saline	Test of phenotype-linked fertility hypothesis	5 µl (1:50 1 mM)	2 µl (2.4 mm)	10 min + 10 min (dark)	500 sperm/male (total number of sperm/spermatophore was also measured –> total number living sperm = %living*total sperm number)	Excluded	Box cox transformation of data, multiple regression models, MANOVA	Klaus et al., 2011
Acheta domesticus (House cricket)	Spermatophore	20 µl	Grace's medium	Effect of sperm competition	1.25 µl 1 mM (1:50 in Grace's medium)	2.5 µl	10 min + 10 min (dark)	Haemocytometer, 5 predetermined areas, 400x magnification, blind, total number of sperm caculated	Not mentioned	ANCOVA	Worthington et al., 2013
Acheta domesticus (House cricket)	Spermatophore	20 µl	Grace's medium	effect of sperm number, ruptured vs. non-ruptured spermatophores, time of spermatophore in solution (0, 5, 10, 20, 30 min)	1.25 µl in 50 µl Grace's medium	2.5 µl	10 min (RT)	Haemocytometer, 5 predetermined areas, 400x magnification, blind, total number of sperm caculated	not mentioned	GLM, ANOVA, ANCOVA	Gress and Kelly, 2011
Anolis sagre (Brown anole)	Ejaculate removed from female's cloaca	20 µl	PBS	effect of paternal and maternal experience and maternal investment on fertility	1.25 µl	2.5 µl	5 min + 5 min	2 chambers of haemocytometer, 400 x magnification, each 5 pre-determined grid squares, blind to treatment, total number of sperm also determined	Not mentioned	Log transformation, GLM	Warner et al., 2013
Apis mellifera (Honey bee)	Ejaculated sperm from endophallus	5 µl (2 µl ejaculate in 200 µl semen dilutent)	"Semen dilutent"	Effect of environmental stress and male age	5 µl	2 µl	10 min + 7 min (dark)	×400 magnifications, we estimated sperm viability twice and counted a minimum of 300 sperm cells per sample, blind to treatment	Not mentioned	Logit transformation, ANOVA	Stürup et al., 2013
Apis mellifera (Honey bee)	Ejaculated sperm from endophallus	300 µl	Kiev buffer	Fresh vs. freeze-thawed samples, effect of mixing, effect of sperm competition	5 µl (1:50 in deionized water)	4 µl	5 min + 0 min (36°C)	Flow cytometry	Not mentioned	Arcsin transformation; ANOVA, Student's *t*-test	Tofilski et al., 2012
Apis mellifera (Honey bee)	Ejaculated sperm from endophallus	150 µl	Modified Kiev sperm diluent	Effect of sublethal miticide concentrations	1 µl (0.1 mM in DMSO)	2 µl (2.4 mM)	10 min +10 min	10 µl aliquot of stained sperm solution, 100X magnification, 5 pictures	Not mentioned	GLMM with binomial error structure and logit link function	Johnson et al., 2013
Apis mellifera (Honey bee)	Male seminal vesicles	200 µl	Collins buffer	Method establishment	1 µl	1 µl	5 min + 5 min (36°C)	Flow cytometry	Not mentioned	Standard deviation, Spearman correlation	Rzymski et al., 2012
Apis mellifera (Honey bee)	Ejaculated from endophallus	5 µl; 300 µl	Semen dilutent	Method establishment	10 µM; 100 nM	400 µM; 12 µM	10 min + 7 min (RT, dark)	Sperm counting of whole sample, flow cytometry	Not mentioned	Pearson product–moment correlation, ANOVA, Wilcoxon signed rank tests, Student's *t*-test	Paynter et al., 2014
Apis mellifera (Honey bee)	Male seminal vesicles, partly and fully everted endophalli, lateral oviducts	150 /500/200/ 300 µl	Kiev buffer	Sperm from different organs, different times of aeration, single vs. multiple drone semen, exposure to pressure	0.75, 1.25, 2.5 µl(1:50 DMSO)	1.25, 2.5, 6 (1:4 dH_2O)	10 min (36°C)	6 aliquots/sample, at least 400 sperm, 400 x magnification	Not mentioned	Arcsin transformation, one-way ANOVA	Gençer et al., 2014
Apis mellifera carnica (Carnolian honey bee)	Ejaculate	300 µl	Kiev buffer	Brood incubation temperature	5 µl (1:50 in sterile ddH_2O)	4 µl	5 min + 0 min (36°C)	Flow cytometry	Not mentioned	Arcsin transformation, nested ANOVA	Czekonska et al., 2013

(Continued)

TABLE 1 | Continued

Species	Organ that sperm was removed from	Amount of sperm suspension mixed with kit	Sperm diluent	Experimental context	Amount/ concentration of SYBR14	Amount/ concentration of PI	Incubation time	Methodological details	Treatment of red and green double- stained cells	Statistical analysis	References
Apis mellifera caucasica (Caucasian honey bee)	Ejaculate and male seminal vesicles	150 µl	modified Kiev buffer	Small drones from laying worker colonies vs. arge drones from queenright colonies	0.6 µl (1:50 in DMSO)	0.6 µl (1:4 in ddH_2O)	10 min (36°C)	6 aliquots/sample, at least 400 sperm	Not mentioned	Square root transformation, Student's *t*-test	Gençer and Kahya, 2011
Atta colombica (leafcutter ant)	Accessory testes	3 µl	Hayes saline	Effect of male glandular fluids vs. Hayes saline	3 µl (2 in 98 Hayes solution)	1 µl	5 min +4 min (dark, RT)	At least 300 sperm, blind measure	Excluded	Generalized Estimating Equations (GEE) with a binomial error distribution and a logit-link function; values relative compared to control (=Hayes)	den Boer et al., 2015
Atta colombica, Acromyrmex echinatior (leafcutter ants)	Accessory testes	2 x 5 µl	Hayes saline solution	Effect of lethal fungal infection in males	5 µl	2 µl	10 min + 7 min (dark)	400x magnification, minimum of 300 sperm/sample, blind to treatment	Not mentioned	Logit transformation, ANCOVA	Stürup et al., 2014
Austropotamobius italicus (crayfish)	Spermatophore	10 µl	Physiological NaCl solution	Relationship between male secondary sexual characters and sperm number, viability, and longevity	3 µl (1:25 in DMSO and then 1 10 in distilled water), 1 mM	1 µl, 2.4 mM	5 min + 5 min	Haemocytometer, over 100 sperm, 2 samples per male for each time point	Not mentioned	Arcsin transformation, mixed model analyses (REML method, degrees of freedom estimated according to the Kenward-Roger method)	Galeotti et al., 2012
Drosophila melanogaster (Fruitfly)	Male seminal vesicles, female seminal receptacle	8 µl	Drosophila Ringer's solution	Test of kit toxicity, effect of female age, sperm senescence and multiple mating	1 µl (1.6 ml of SYBR-14 in 10 ml of DMSO)	1 µl (1.6 ml in 18 ml of Drosophila Ringer's solution)	3 min + 1 min (dark)	400 x, pictures of all sperm, blind to treatment	Counted as dead	Arcsine square root transformation, ANCOVA	Radhakrishnan and Fedorka, 2011
Drosophila melanogaster (Fruitfly)	Male seminal vesicle	2.5 µl	HEPES-buffered saline	Effect of male age classes and naturally occurring genotypes	1.25 µl of 2 µl of SYBR-14 and 44 µl of PI per 1 ml of HEPES-buffered saline	s. SYBR14	5 min	200 x, total sperm sample	Counted as dead	GLMs with binomial errors and logit link	(Decanini et al., 2013)
Gryllus veletis (Spring field cricket)	Spermatophore (removed immediately after mating)	5 µl	Beadle saline	Test of phenotype-linked fertility hypothesis	5 µl (1:50, 1 mm)	2 µl (2.4 mm)	10 min (dark, RT)	Sperm number assessed, images taken immediately, at least 100 sperm/male	Excluded	GLM	Fitzsimmons and Bertram, 2013
Hyla versicolor (Gray tree frog)	Testis	25 µl	Amphibian's Ringer solution	Test of phenotype-linked fertility hypothesis	0.5 µl, 40 nmol/l	0.5 µl, 46 µmol/l	5 min + 5 min (RT)	200× magnification, 2 independent samples for each male; total amount of sperm quantified	Not mentioned	Log- or arcsine-Square-root transformation PROC MIXED	Doyle, 2011

(Continued)

TABLE 1 | Continued

Species	Organ that sperm was removed from	Amount of sperm suspension mixed with kit	Sperm diluent	Experimental context	Amount/ concentration of SYBR14	Amount/ concentration of PI	Incubation time	Methodological details	Treatment of red and green double- stained cells	Statistical analysis	References
Macrolophus pygmaeu, Nesidiocoris tenuis (Mirid bugs)	Female receive and storage organs	5 µl	Beadle's saline	Species differences	0.5 µl 1:50 1 mM	2 µl (2.4 mM)	10 min (23°C, dark) + 0 min	13 replicates per species	Not mentioned	one-Way ANOVAs after normalization by log(x + 1) (counts) or arcsin transformation (percentages) Tukey test	Franco et al., 2011
Mnais pruinosa, Calopteryx cornelia (damselflies)	Male seminal vesicles, female bursa copulatrix, spermatheca	10 µl	Grace's insect cell culture medium	Effect of mating stages, male and female sperm storage organs	1 µl of a 0.004 mmol/l (in DMSO)	10 µl 0.048 mmol/L (in 50 Grace's)	5 min (dark, RT)	8 µl of solution, measured within 20 min, mean = 84 sperm/sample, 200x magnification	Counted as viable	Arcsin transformation, *t*-tests, paired sample *t*-tests, ANOVA, correlation,	Tsuchiya and Hayashi, 2010
Poecilia reticulate (guppy)	Ejaculate	10 µl	150 mM KCl-solution	Effect of ovarian fluidand male sexual ornamentation on sperm viability and survival	10 µl (1:50 1 mM)	2 µl 2.4 mM	10 min + 10 min (dark)	ca. 100 sperm/trial, alternating order between control and treatment	Not mentioned	Arcsin square-root transformation, paired *t*-tests, linear mixed-effects model	Gasparini and Evans, 2013
Scaptotrigona aff. Depilis (stingless bee)	Seminal vesicle	? µl	Hayes saline	Young vs. old males, protocol development (pH, amount of saline)	? µl	? µl	?	First 400 cells from center of cover slip	Not mentioned	Binomial GLMs	Meneses et al., 2014
Social ants and bees	Accessory testes	Depending on species	Hayes saline	Effect of own vs. foreign seminal fluid in monogamous vs. polygamous species	?	?	?	At least 300 randomly selected spermatozoa	Excluded	Log transformation, regression analysis, χ^2-test	den Boer et al., 2010
Teleogryllus oceanicus (Australian field cricket)	Spermatophore	5 µl	Beadle's saline	Trade-offs between sperm quality and immunity	5 µl (1:50 1 mM)	2 µl	5 min + 5 min (dark)	First 500 sperm observed on the slide was scored across multiple randomly-positioned fields on the slide	Counted as dead	Arcsin transformation, GLM	Dowling and Simmons, 2012
Xiphophorus nigrensis (Panuco swordtail)	Ejaculate	9.5 µl	"Sperm extender"	Sneaking vs. courting males, before and 3 h after sperm activation	0.5 µl (200 nm)	1 µl (12 µm)	10 min + 10 min (RT)	Pictures at 100x, evaluated by ImageJ	Not mentioned	GLM with a logit link and scaled dispersion parameter	Juenger et al., 2011

PBS, phosphate-buffered saline; PI, propidium iodide, red fluorescent dye that stains nucleic acids of sperm with damaged membrane red; RT, room temperature; SYBR 14, fluorescent dye that stains nucleic acids of sperm with intact membrane green. ?, No information provided.

TABLE 2 | Pearson correlation of total number of sperm and proportion of live sperm (sperm viability, SV) in *Cimex* and *Drosophila* for the different experiments.

Experiment	Species	*t*	d.f.	*p*
Non-blind, standard (30 min)	*C. lectularius*	1.54	26	0.14
Non-blind, standard (0 min)	*C. lectularius*	0.96	26	0.34
Blind, standard (30 min)	*C. lectularius*	−1.03	26	0.20
Blind, standard (0 min)	*C. lectularius*	−0.46	26	0.65
Blind, time series (0 min)	*C. lectularius*	−1.30	26	0.20
Blind, time series (30 min)	*C. lectularius*	−0.8	26	0.44
Blind, standard (0 min)	*D. melanogaster*	1.85	13	0.09
Blind, standard (30 min)	*D. melanogaster*	0.20	13	0.84
Blind, time series (0 min)	*D. melanogaster*	1.00	8	0.35
Blind, time series (30 min)	*D. melanogaster*	1.16	8	0.28

For the total number of sperm, see Supplementary Figure 1.

TABLE 3 | Location, time, and interaction effects on sperm viability, determined by a generalized linear mixed model for the non-blind standard experiment with *Cimex lectularius* (area highlighted in gray, $n = 14$).

Factor	Estimate	z-value	*p*-value	2.5% CI	97.5% CI
Intercept	0.01 (±0.41)	0.03	0.97	−0.8	0.82
Location	1.34 (±0.48)	2.81	0.005	0.37	2.28
Time	0.20 (±0.48)	0.41	0.68	−0.78	1.16
Location*time	−0.46 (±0.64)	−0.72	0.47	−1.74	0.84
Intercept	1.90 (±0.30)	6.31	0.0001	1.3	2.51
Location	0.46 (±0.44)	1.06	0.29	−0.41	2.51
Time	−0.94 (±0.09)	−10.58	< 0.0001	−1.12	−0.77
Location*time	−0.32 (±0.13)	−2.42	0.02	−0.59	−0.06

Location is either cranial or caudal, time either t0 or t15. Location, time and interaction effects on sperm viability, determined by a generalized linear mixed model for the blind time series experiment with Cimex lectularius (white area, n = 10). Location is either cranial or caudal, time either t_0, t_5, t_{15} or t_{30}.

independent of its dye and therefore, did not invalidate the cross-sectional results. While both stress test and longitudinal approach represent a complication of the SV method, both have a number of advantages, which we discuss below. We suggest that future researchers may incorporate this test into their method portfolio if, for example, analyzing differences in sperm quality across species, across male and female sperm storage organs, across male or female reproductive fluids (Scaggiante et al., 1999; Rosengrave et al., 2008; Simmons et al., 2009; Doyle, 2011; Otti et al., 2013) or across male age cohorts. This may be, particularly useful when more sophisticated equipment is not at hand or impossible to use.

SV Heterogeneity

We found several sources of heterogeneity of SV staining in ecology and evolution research. First, the protocols by different researchers differed in various details such as buffer, incubation time, dye concentration. This unsatisfactory situation may be caused by large within-sample heterogeneity, i.e., a lack of significant between-image repeatability, which we confirmed empirically for our two study species. We believe that this heterogeneity has a biological cause because the procedural precision of the counting itself was very high. Because we restricted our analysis to studies citing Holman (2009), it is possible that studies that did not cite this author may have followed procedures that are even less stringent, including our own (Otti et al., 2013). We also note that most of these studies concerned invertebrates.

Second, we found another intrinsic difference between our two species. *Cimex* sperm was visible immediately after staining whereas the visibility of *Drosophila* sperm after staining was substantially delayed—even 30 min after staining, it was not clear whether all sperm had been stained. Studies by researchers that are either unaware of this difference, or had developed protocols to circumvent the delayed staining will, therefore, differ in various protocol details. It is important to consider that such differences may in theory also occur within a species, when sperm is harvested from different organs. We hope that our time series approach may be useful to assess this variation. Our time series protocol examined the same sperm sample repeatedly and therefore allowed us to establish that the visibility was not biased with respect to dead (red) or live (green) sperm. This longitudinal approach appears to us an important way of characterizing sperm quality because it allows the tracking of individual sperm. In theory, total visible sperm number can increase with time (as in our study) or decrease, such as if photobleaching would occur.

Third, both across studies as well as within our data set there was variation with respect to density-dependent SV, which either does not exist (or has not been tested for) or which is positive, i.e., SV is higher at higher sperm density. It is possible that density as a source of variation in SV has a biological basis, the so-called respiratory-dilution effect, where high dilution increases stress via increased endogenous respiration (Mann, 1967). Importantly, even in the case that SV and sperm density would be positively correlated (**Table 2**), the sperm number differences across our experimental protocols did not cause initial differences in SV (**Table 6**).

Fourth, we found variation across studies in measuring SV blind and we examined this issue empirically. We found that whether or not the SV measurement is carried out blind had a strong effect on SV. In our case, results that supported a specific hypothesis were confirmed. Given that few studies practice this widely-recommended procedure, we can only reinforce all previous calls and suggest an urgent need to implement blind SV measurement. Where exactly observer bias arose during SV measurements is not currently known but we believe that an unconscious biased choice of counting areas plays an important role. Our blind/non-blind comparison involved some further protocol adjustments or standardization (e.g., male age, dissection buffer). However, these adjustments seem small, such as the brief period that the insects were situated in dissection buffer or, as in the case of male age, are predicted to increase the effects, and hence we would err on the wrong side. We, therefore, believe that the blind/non-blind difference is the major protocol aspect responsible for the results observed.

Protocol Standardization and Recommendations

In addition to the recommended blind measurement, the standardization in protocols in measuring a character that is

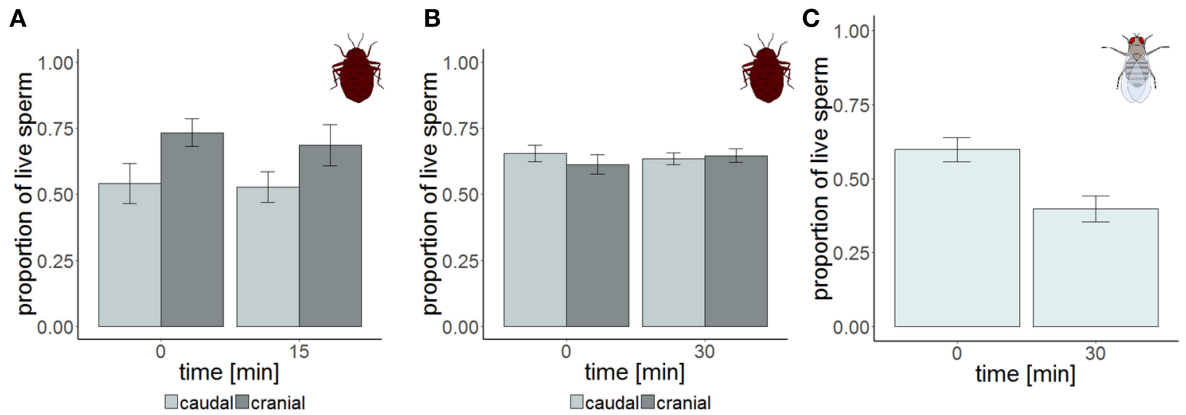

FIGURE 2 | Sperm viability in the cranial (dark gray) and the caudal (light gray) part of the seminal vesicle of *Cimex lectularius* at two different points in time. The samples were measured either non-blind (**A**, $n = 14$) and supported the hypothesis that sperm should have higher viability at the cranial site, or were blind (**B**, $n = 15$). Sperm viability in the seminal vesicle of *Drosophila melanogaster* (**C**, $n = 15$). Bars show mean ± s.e.

TABLE 4 | Intra-class correlation coefficient (ICC) of sperm viability across four pictures of the same subsample in a non-blind standard count in *Cimex lectularius* (area highlighted in light gray, $n = 14$), a blind count *Cimex lectularius* (area highlighted in dark gray, $n = 15$), and a blind count in *Drosophila melanogaster* (white area, $n = 15$).

Species	Experiment	Time	Location	ICC	2.5% CI	97.5% CI
C. lectularius	Non-blind, standard	0	Cranial	0.45	0.18	0.73
			Caudal	0.59	0.34	0.82
		15	Cranial	0.74	0.53	0.89
			Caudal	0.34	0.09	0.65
C. lectularius	Blind, standard	0	Cranial	0.66	0.41	0.86
			Caudal	0.59	0.32	0.82
		30	Cranial	0.32	0.06	0.64
			Caudal	0	−0.19	0.26
D. melanogaster	Blind, standard	0		0.47	0.2	0.74
		30		0.42	0.16	0.7

as severely environment-dependent as SV, is a precondition to compare SV across studies. For example, we found no effect of time on SV in bedbugs. Surprisingly, in cross sections, SV sometimes was higher at t_{15} or t_{30}, than at t_0 (note that in *Cimex* sperm visibility did not change over time—**Figure 4**). As obviously dead sperm cannot revive, the distribution of live and dead sperm in the seminal vesicle of *Cimex* and *Drosophila* was heterogeneous and mixing by pipetting insufficient to homogenize sperm. Consequently, different subsamples vary in their initial proportion of viable sperm, which is also reflected in the low repeatability of SV. Sometimes, as in *Cimex* (but not *Drosophila*) in our study, the low repeatability can mask biological effects. Unlike current practice (**Table 1**), researchers may wish to incorporate the analysis of heterogeneity in the lab routine, and report the outcomes. The repeatability pattern may be identified as species-specific only when the same buffer, incubation duration and concentrations are used. In none of our experiments was the total number of sperm correlated with SV, in this case contradicting (Holman, 2009) but we support that study's cautionary remarks because again protocols and species may differ between in the sperm density-sperm mortality relationship. While our emphasis was to present the temporal variation of SV, this will of course not release researchers from identifying the optimal dye concentrations and incubation times for SV measurement in their study species. And our study may provide a particularly suitable, though initially unintended, example with respect to incubation time. While in bedbugs no incubation was necessary, in *Drosophila*, the protocol would need refinement. Methodologically, *Drosophila* sperm would require incubation times, that would biologically be prohibitively long. We recommend to aim at dilution protocols that keep sperm densities low. However, our method allowed us to analyse the visibility of sperm over time and we found it equally increased for red and green sperm. As such, even the lack of incubation time, as in our four-image cross sampling, appears an adequate reflection of SV. In any case, incubation time should be strictly standardized. We note that flow cytometry-based counting methods also require the detection of stained sperm, and therefore, may also need to account for possible stainability differences.

Another precondition for comparing data across studies is the correct statistical analysis. Estimating the effect size errors made by inadequate statistical treatment was beyond the scope of our paper but we reinforce calls (Holman, 2009) for the correct statistical analysis and point to binomial GLMMs that are recommended for analyzing proportional data (Warton and Hui, 2011). Especially the widely applied arcsine square root transformation (**Table 1**) should not be used for binomial data as the transformation decreases the power and interpretability of the model (Warton and Hui, 2011).

Sperm Viability vs. Sperm Quality

SV *per se* may, or may not be informative about the function or success of the living sperm in that sample (see Introduction).

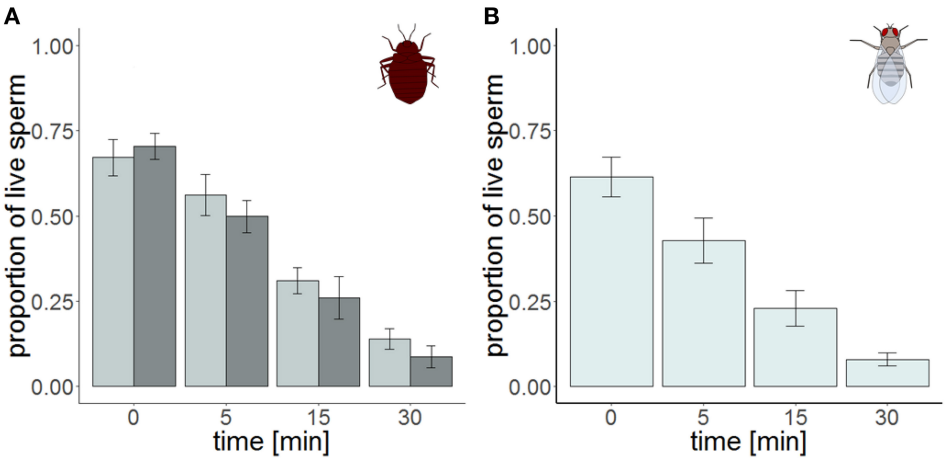

FIGURE 3 | Sperm viability in the cranial (dark gray) and the caudal (light gray) part of the seminal vesicle of *Cimex lectularius* (**A**, $n = 10$) and in seminal vesicle of *Drosophila melanogaster* (**B**, $n = 10$) at four different points in time. Bars show mean ± s.e.

TABLE 5 | Time effect on sperm viability, determined by a generalized linear mixed model for the blind standard experiment in *Drosophila melanogaster* (area highlighted in gray, $n = 15$).

Factor	Estimate	z-value	*p*-value	2.5% CI	97.5% CI
Intercept	0.43 (±0.17)	2.46	0.01	0.09	0.78
Time	−0.03 (±0.008)	−3.1	0.002	−0.04	−0.009
Intercept	0.32 (±0.28)	1.18	0.24	−0.27	0.91
Time	−0.12 (±0.007)	−15.39	<0.0001	−0.12	−0.09

Time is either t_0 or t_{30}. Time effect on sperm viability, determined by a generalized linear mixed model for the blind time series experiment in Drosophila melanogaster (white area, $n = 10$). Time is either t_0, t_5, t_{15}, or t_{30}.

We presented a method to repeatedly measure the same sample in a stressor medium (sperm stress test). Arguably, the resulting mortality rate of the living sperm between time x and x+1 is a better indicator of the performance or membrane properties of living sperm, and therefore of sperm quality, than the number of dead sperm at time x. In principle, the mortality rate could be calculated from cross sectional samples at different time points (see Otti et al., 2013 for an application) but this method only worked in one of the two species we looked at (fruit flies), surprisingly not for bedbugs, the species that Otti et al. (2013) were using, and for which in our study the cross sections were too variable to provide meaningful slope estimates for sperm mortality. We note that the stress test bears similarities to measuring sperm swimming speed over time and we suggest it is likely to have similar pitfalls (see Reinhardt and Otti, 2012 for a discussion of these). For example, trade-offs may exist between SV at t_0 and the slope of the decline toward t_{30}.

However, our proposed stress test resulted in improved precision and allowed in one species the uncovering of a biological process that was hidden in cross section sampling—sperm stratification. A potential disadvantage of the time series sampling is its longer duration because no other samples can be processed in between or requires highly standardized adjustment of the xy microscopy table. It is also important to note that the stress to sperm is not necessarily reflecting the exact natural response to an environmental factor because the osmotic stress from the distilled water is perhaps aggravated by the toxicity of the SV staining kit, oxygen, or photostress from the excitation source.

Sperm Stratification

Although representing a side result, we wish to briefly comment on intra-male stratification of sperm quality. We found that sperm from the cranial part died faster compared to sperm from the caudal part. This pattern was opposite to what was predicted if aged sperm accumulate toward the ejaculation site (Reinhardt, 2007). As an *a posteriori* explanation, we could think of the possibility that frequent sperm aging might select for on an optimal sperm age distribution at the site of ejaculation at the evolutionary average mating. For example, sperm may be released from the testis but may mature to full function only while moving toward the ejaculation site. Regardless, our results suggest that the repeated matings of a male may not involve identical sperm qualities in species in which sperm stratification occurs. This fact seems important when drawing conclusion about the genetic quality of a male from sperm competition results.

CONCLUSION

We presented a sperm stress test as a more meaningful method than SV to assess future sperm quality. We strongly recommend a blind selection of the sperm count areas when sperm quality is to be measured by SV staining. This will be important in species with long and/or clumped spermatozoa (like *D. melanogaster* and *C. lectularius*) that make the largely unbiased flow cytometry (Holman, 2009) impossible. Just as measuring the

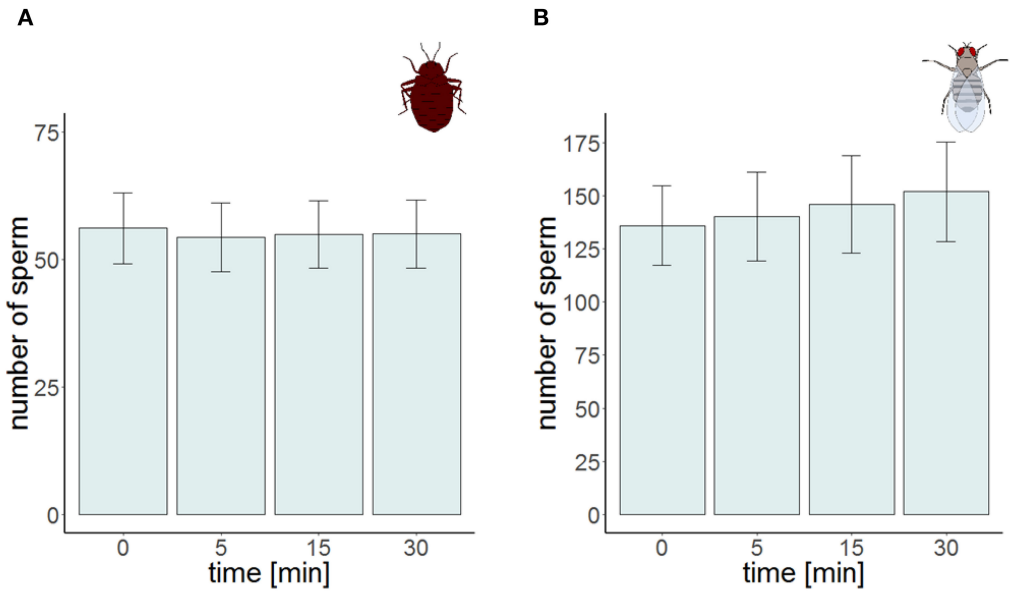

FIGURE 4 | Sperm number at the four different points in time for the time series experiment in *Cimex lectularius* (**A**, $n = 10$) and in *Drosophila melanogaster* (**B**, $n = 10$). Bars show mean ± s.e.

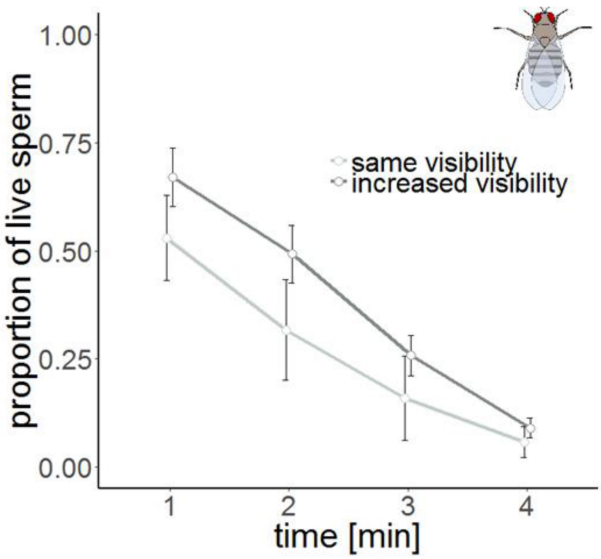

FIGURE 5 | Sperm viability in samples with the same visibility ($n = 4$, light gray) and increased visibility ($n = 6$, dark gray) in *Drosophila melanogaster* at four different points in time. Bars show mean ± s.e.

TABLE 6 | Sperm numbers and viability for the different protocols in *Cimex lectularius* and *Drosophila melanogaster*.

Experiment	Species	Mean number of sperm counted per sample	Mean proportion (%) of live sperm at t_0
Non-blind, standard	*C. lectularius*	374.0 (22–2095)	65.43% (15.70–95.00%)
Blind, standard	*C. lectularius*	326.4 (79–803)	63.50% (26.75–83.33%)
Blind, time series	*C. lectularius*	55 (5–133)	68.76% (40.00–88.57%)
Blind, standard	*D. melanogaster*	427.6 (199–797)	59.83% (28.33–82.60%)
Blind, time series	*D. melanogaster*	145.8 (41–238)	61.49% (37.14–89.33%)

Data are mean (minimum–maximum).

repeatability is an important approach to estimate the consistency of phenotypes (Nakagawa and Schielzeth, 2010), estimating the homogeneity of SV within sperm samples will be important. If SV is homogenous the standard blind stress test will be sufficient where several pictures represent the entire ejaculate. If SV is heterogeneous, researchers might try whether the here proposed time series measurement gives less variable results. This time series measurement will additionally provide information about the visibility of stained sperm which can depend on species and sperm density but which show was unlikely to alter our results. We point out that the mortality of many other cells is assessed by viability staining kits and many of these studies might benefit from our proposed assay of cell quality using cell mortality, rather than viability *per se*.

AUTHOR CONTRIBUTIONS

BE and KR conceived the study, wrote the manuscript; BE and RG carried out the various SV protocols; BE carried out the literature review, analyzed the data. All authors critically read, and approved, its final version.

FUNDING

The work is supported by the Zukunftskonzept of the Technische Universität Dresden awarded by the Deutsche Forschungsgemeinschaft through the Excellence Initiative.

ACKNOWLEDGMENTS

We thank Christin Froschauer, Cornelia Wetzker, Oliver Heinzel, and Wei Dong from the Applied Zoology lab in Dresden for helping with the blinded and non-blinded experiments. We thank Biz Turnell, Cornelia Wetzker, and Ralph Dobler from the Applied Zoology lab in Dresden and Oliver Otti from the Tierökologie I lab in Bayreuth for reading and discussing various drafts of the manuscript and the Applied Zoology lab for feedback. We further thank the reviewers for their valuable comments that helped to improve and clarify this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2017.00165/full#supplementary-material

REFERENCES

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:*14065823. doi: 10.18637/jss.v067.i01

Browne, W. J., Subramanian, S. V., Jones, K., and Goldstein, H. (2005). Variance partitioning in multilevel logistic models that exhibit overdispersion. *J. R. Stat. Soc. Ser. A* 168, 599–613. doi: 10.1111/j.1467-985X.2004.00365.x

Czekonska, K., Chuda-Mickiewicz, B., and Chorbinski, P. (2013). The effect of brood incubation temperature on the reproductive value of honey bee (*Apis mellifera*) drones. *J. Apic. Res.* 52, 96–105. doi: 10.3896/IBRA.1.52.2.19

Decanini, D. P., Wong, B. B., and Dowling, D. K. (2013). Context-dependent expression of sperm quality in the fruitfly. *Biol. Lett.* 9:20130736. doi: 10.1098/rsbl.2013.0736

den Boer, S. P. A., Stürup, M., Boomsma, J. J., and Baer, B. (2015). The ejaculatory biology of leafcutter ants. *J. Insect Physiol.* 74, 56–62. doi: 10.1016/j.jinsphys.2015.02.006

den Boer, S. P., Baer, B., and Boomsma, J. J. (2010). Seminal fluid mediates ejaculate competition in social insects. *Science* 327, 1506–1509. doi: 10.1126/science.1184709

Dowling, D. K., and Simmons, L. W. (2012). Ejaculate economics: testing the effects of male sexual history on the trade-off between sperm and immune function in Australian crickets. *PLoS ONE* 7:e30172. doi: 10.1371/journal.pone.0030172

Doyle, J. M. (2011). Sperm depletion and a test of the phenotype-linked fertility hypothesis in gray treefrogs (*Hyla versicolor*). *Can. J. Zool.* 89, 853–858. doi: 10.1139/z11-060

Fitzsimmons, L. P., and Bertram, S. M. (2013). No relationship between long-distance acoustic mate attraction signals and male fertility or female preference in spring field crickets. *Behav. Ecol. Sociobiol.* 67, 885–893. doi: 10.1007/s00265-013-1511-z

Forstmeier, W., Wagenmakers, E. J., and Parker, T. H. (2016). Detecting and avoiding likely false-positive findings–a practical guide. *Biol. Rev.* 92, 1941–1968 doi: 10.1111/brv.12315

Franco, K., Jauset, A., and Castañé, C. (2011). Monogamy and polygamy in two species of mirid bugs: a functional-based approach. *J. Insect Physiol.* 57, 307–315. doi: 10.1016/j.jinsphys.2010.11.020

Galeotti, P., Bernini, G., Locatello, L., Sacchi, R., Fasola, M., and Rubolini, D. (2012). Sperm traits negatively covary with size and asymmetry of a secondary sexual trait in a freshwater crayfish. *PLoS ONE* 7:e43771. doi: 10.1371/journal.pone.0043771

Garcia-González, F., and Simmons, L. W. (2005). Sperm viability matters in insect sperm competition. *Curr. Biol.* 15, 271–275. doi: 10.1016/j.cub.2005.01.032

Gasparini, C., and Evans, J. P. (2013). Ovarian fluid mediates the temporal decline in sperm viability in a fish with sperm storage. *PLoS ONE* 8:e64431. doi: 10.1371/journal.pone.0064431

Gençer, H. V., and Kahya, Y. (2011). Are sperm traits of drones (*Apis mellifera* L.) from laying worker colonies noteworthy? *J. Apic. Res.* 50, 130–137. doi: 10.3896/IBRA.1.50.2.04

Gençer, H. V., Kahya, Y., and Woyke, J. (2014). Why the viability of spermatozoa diminishes in the honeybee (*Apis mellifera*) within short time during natural mating and preparation for instrumental insemination. *Apidologie* 45, 757–770. doi: 10.1007/s13592-014-0295-0

Gress, B. E., and Kelly, C. D. (2011). Is sperm viability independent of ejaculate size in the house cricket (*Acheta domesticus*)? *Can. J. Zool.* 89, 1231–1236. doi: 10.1139/z11-103

Holman, L. (2009). Sperm viability staining in ecology and evolution: potential pitfalls. *Behav. Ecol. Sociobiol.* 63, 1679–1688. doi: 10.1007/s00265-009-0816-4

Holman, L., Head, M. L., Lanfear, R., and Jennions, M. D. (2015). Evidence of experimental bias in the life sciences: why we need blind data recording. *PLoS Biol.* 13:e1002190. doi: 10.1371/journal.pbio.1002190

Hunter, F., and Birkhead, T. (2002). Sperm viability and sperm competition in insects. *Curr. Biol.* 12, 121–123. doi: 10.1016/S0960-9822(01)00647-9

Johnson, R. M., Dahlgren, L., Siegfried, B. D., and Ellis, M. D. (2013). Effect of in-hive miticides on drone honey bee survival and sperm viability. *J. Apic. Res.* 52, 88–95. doi: 10.3896/IBRA.1.52.2.18

Juenger, T., Bolnick, D., and Rosenthal, G. (2011). *Sperm Competition and the Evolution of Alternative Reproductive Tactics in the Swordtail Xiphophorus nigrensis (Poeciliidae)*. Doctoral dissertation, The University of texas at Austin.

Klaus, S. P., Fitzsimmons, L. P., Pitcher, T. E., and Bertram, S. M. (2011). Song and sperm in crickets: a trade-off between pre-and post-copulatory traits or phenotype-linked fertility? *Ethology* 117, 154–162. doi: 10.1111/j.1439-0310.2010.01857.x

Mann, T. (1967). Sperm metabolism. *Fertil. Comp. Morphol. Biochem. Immunol.* 1, 99–116.

Meneses, H. M., Koffler, S., Freitas, B. M., Imperatriz-Fonseca, V. L., and Jaffé, R. (2014). Assessing sperm quality in stingless bees. *Sociobiology* 61, 517–522. doi: 10.13102/sociobiology.v61i4.517-522

Milinski, M. (1997). How to avoid seven deadly sins in the study. *Adv. Stud. Behav.* 26, 159–180. doi: 10.1016/S0065-3454(08)60379-4

Morrell, J., and Rodriguez-Martinez, H. (2009). Biomimetic techniques for improving sperm quality in animal breeding: a review. *Open Androl. J.* 1, 1–9. doi: 10.2174/1876827X00901010001

Nakagawa, S., and Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol. Rev.* 85, 935–956. doi: 10.1111/j.1469-185X.2010.00141.x

Otti, O., McTighe, A. P., and Reinhardt, K. (2013). *In vitro* antimicrobial sperm protection by an ejaculate-like substance. *Funct. Ecol.* 27, 219–226. doi: 10.1111/1365-2435.12025

Parker, G. A. (1970). Sperm competition and its evolutionary consequences in the insects. *Biol. Rev.* 45, 525–567. doi: 10.1111/j.1469-185X.1970.tb01176.x

Parker, G. A., and Pizzari, T. (2010). Sperm competition and ejaculate economics. *Biol. Rev.* 85, 897–934. doi: 10.1111/j.1469-185X.2010.00140.x

Paynter, E., Baer-Imhoof, B., Linden, M., Lee-Pullen, T., Heel, K., Rigby, P., et al. (2014). Flow cytometry as a rapid and reliable method to quantify sperm viability in the honeybee *Apis mellifera*. *Cytometry Part A* 85, 463–472. doi: 10.1002/cyto.a.22462

Pitnick, S., Hosken, D. J., and Birkhead, T. R. (2009). *Sperm Biology: An Evolutionary Perspective, Chapter 3: Sperm Morphological Diversity, 1st Edn*. Burlington, ON; San Diego, CA; Oxford: Academic press.

R Development Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: Foundation for Statistical Computing.

Radhakrishnan, P., and Fedorka, K. M. (2011). Influence of female age, sperm senescence and multiple mating on sperm viability in female *Drosophila melanogaster*. *J. Insect Physiol.* 57, 778–783. doi: 10.1016/j.jinsphys.2011.02.017

Reinhardt, K. (2007). Evolutionary consequences of sperm cell aging. *Q. Rev. Biol.* 82, 375–393. doi: 10.1086/522811

Reinhardt, K., and Otti, O. (2012). Comparing sperm swimming speed. *Evol. Ecol. Res.* 14, 1039–1056.

Reinhardt, K., Dobler, R., and Abbott, J. (2015). An ecology of sperm: sperm diversification by natural selection. *Annu. Rev. Ecol. Evol. Syst.* 46, 435–459. doi: 10.1146/annurev-ecolsys-120213-091611

Reinhardt, K., Naylor, R., and Siva–Jothy, M. T. (2003). Reducing a cost of traumatic insemination: female bedbugs evolve a unique organ. *Proc. R. Soc. Lond. B Biol. Sci.* 270, 2371–2375. doi: 10.1098/rspb.2003.2515

Rosengrave, P., Gemmell, N. J., Metcalf, V., McBride, K., and Montgomerie, R. (2008). A mechanism for cryptic female choice in chinook salmon. *Behav. Ecol.* 19, 1179–1185. doi: 10.1093/beheco/arn089

Ryne, C. (2009). Homosexual interactions in bed bugs: alarm pheromones as male recognition signals. *Anim. Behav.* 78, 1471–1475. doi: 10.1016/j.anbehav.2009.09.033

Rzymski, P., Langowska, A., Fliszkiewicz, M., Poniedziałek, B., Karczewski, J., and Wiktorowicz, K. (2012). Flow cytometry as an estimation tool for honey bee sperm viability. *Theriogenology* 77, 1642–1647. doi: 10.1016/j.theriogenology.2011.12.009

Scaggiante, M., Mazzoldi, C., Petersen, C., and Rasotto, M. (1999). Sperm competition and mode of fertilization in the grass goby *Zosterisessor ophiocephalus* (Teleostei: Gobiidae). *J. Exp. Zool.* 283, 81–90. doi: 10.1002/(SICI)1097-010X(19990101)283:1<81::AID-JEZ9>3.0.CO;2-9

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089

Simmons, L., Roberts, J., and Dziminski, M. (2009). Egg jelly influences sperm motility in the externally fertilizing frog, *Crinia georgiana*. *J. Evol. Biol.* 22, 225–229. doi: 10.1111/j.1420-9101.2008.01628.x

Stürup, M., Baer, B., and Boomsma, J. J. (2014). Short independent lives and selection for maximal sperm survival make investment in immune defences unprofitable for leaf-cutting ant males. *Behav. Ecol. Sociobiol.* 68, 947–955. doi: 10.1007/s00265-014-1707-x

Stürup, M., Baer-Imhoof, B., Nash, D. R., Boomsma, J. J., and Baer, B. (2013). When every sperm counts: factors affecting male fertility in the honeybee *Apis mellifera*. *Behav. Ecol.* 24, 1192–1198. doi: 10.1093/beheco/art049

Tofilski, A., Chuda-Mickiewicz, B., Czekonska, K., and Chorbinski, P. (2012). Flow cytometry evidence about sperm competition in honey bee (*Apis mellifera*). *Apidologie* 43, 63–70. doi: 10.1007/s13592-011-0089-6

Tsuchiya, K., and Hayashi, F. (2010). Factors affecting sperm quality before and after mating of calopterygid damselflies. *PLoS ONE* 5:e9904. doi: 10.1371/journal.pone.0009904

Warner, D. A., Kelly, C. D., and Lovern, M. B. (2013). Experience affects mating behavior, but does not impact parental reproductive allocation in a lizard. *Behav. Ecol. Sociobiol.* 67, 973–983. doi: 10.1007/s00265-013-1523-8

Warton, D. I., and Hui, F. K. (2011). The arcsine is asinine: the analysis of proportions in ecology. *Ecology* 92, 3–10. doi: 10.1890/10-0340.1

Wolak, M., and Wolak, M. M. (2015). *Package "ICC". Facilitating Estimation of the Intraclass Correlation Coefficient*.

World Health Organization (2010). *WHO Laboratory Manual for the Examination and Processing of Human Semen*, *5th Edn*. Geneva.

Worthington, A. M., Gress, B. E., Neyer, A. A., and Kelly, C. D. (2013). Do male crickets strategically adjust the number and viability of their sperm under sperm competition? *Anim. Behav.* 86, 55–60. doi: 10.1016/j.anbehav.2013.04.010

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer, FGG, and handling Editor declared their shared affiliation.

Copyright © 2017 Eckel, Guo and Reinhardt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.