The results of lithic experiments performed on glass cores are applicable to other raw materials

Tamara Dogandžić1,2 · Aylar Abdolazadeh2 · George Leader2,3,4 · Li Li2,5 · Shannon P. McPherron1 · Claudio Tennie5 · Harold L. Dibble2

Received: 29 April 2019 / Accepted: 2 December 2019 © The Author(s) 2020

Abstract
About 10 years ago, a new experimental design, based on a mechanical flaking apparatus, allowed complete control over several independent variables essential to flintknapping. This experimental setting permitted the investigation of more fundamental aspects of stone technology, including the effect of particular platform attributes, core surface morphology, and the application of force on flake size and shape. These experiments used cores made of glass that were molded to exact configurations. Here we set out to investigate whether results obtained from experiments on glass cores can be extended to other materials, in this case varieties of basalt, flint, and obsidian that were cut to the exact core configurations. We focused on the relationships between the independent variables of exterior platform angle and platform depth and dependent variables of overall size (weight or mass), volume, and linear dimensions. It was found that in almost every comparison, all four materials show similar relationships in nature and degree. What differs instead is the amount of force needed to detach a flake. In other words, given the same core morphology and platform attributes the resulting flakes will be the same, but harder materials require more force to remove the flake. These results were additionally verified on Middle Paleolithic archeological materials made mostly on Late Cretaceous flints. Our results demonstrate that experiments using glass cores are valid and can be generalized and extended to other materials.

Keywords Controlled experiments · Raw materials · Glass cores · Exterior platform angle · Platform depth

Introduction
Controlled experiments on flake manufacture have a long history in lithic studies (Bonnichsen 1977; Dibble and Whittaker 1981; Faulkner 1972; Pelcin and Dibble 1995; Pelcin 1997a, b, 1998; Speth 1972, 1975, 1981). These experiments were designed to investigate factors that influence size and shape of a flake. The relationship between platform variables and flake size was further used to estimate the original size of the blank before it has been reduced by retouch. Reduction intensity can be quantified by comparing the original mass and the mass of a retouched element (Dibble 1987, 1995). This aspect has been valuable for curation studies, particularly for the questions of the Middle Paleolithic industrial variability since resharpening events drive much of the typological variability as established by F. Bordes (Bordes 1961; Dibble 1987, 1995; Hiscock and Clarkson 2007, 2015; Turq 1989). Tight correlation of platform attributes and flake size has motivated a number of studies to refine platform measurements in order to improve this relationship and further address reduction intensity in lithic assemblages (Braun et al. 2008; Clarkson and Hiscock 2011; Davis and Shea 1998; Muller and Clarkson 2014; Shott et al. 2000).

One of the primary criticisms of controlled experiments has been about questions of validity. The criticism was that the products themselves, and/or the means by which force was...
applied to remove them, bear little resemblance to knapped flakes and cores. About 10 years ago, a new experimental design was developed with the intention of enhancing the validity of the experimental flaking process and the resulting flakes themselves (Dibble and Rezek 2009). While the resulting cores and flakes more closely resembled knapped artifacts, the cores themselves were made of soda-lime glass (henceforth glass), reliably molded to specific shapes. This allowed for experiments on the role of various independent variables (those under the control of the knapper) on specific flake characteristics. Cores produced from the same mold were rather identical in size and shape, thus allowing important variables to be held constant—a major advantage in any experimentation. Different core designs could be produced to also examine the effects of core surface morphology on flake variation (Rezek et al. 2011).

Fig. 1 Examples of cores made of different raw materials (upper left to lower right: flint, basalt, obsidian, glass). The striking surface of the cores shows some slight variability due to hand production.

Over the course of numerous glass-core experiments, several independent variables were examined, including exterior platform angle, platform depth, hammer speed, force, angle of blow (Pelcin and Dibble 1995), core morphology (Rezek et al. 2011), hammer type (Magnani et al. 2014), and platform beveling (Leader et al. 2017). The results of these experiments have helped to describe and quantify how knappers achieve particular results, with direct implications for interpreting the archeological record (Lin et al. 2013; Režek et al. 2018). While glass itself is highly amenable to percussion flaking, it still leaves the question as to the applicability and the validity of the results of these experiments to the kinds of raw materials actually knapped in the past.

Much of the archeological literature dealing with the role that raw materials play in variation among lithic assemblages deals with aspects of accessibility and quantity (Andrefsky...
Materials and methods

In this experiment, we examined if the effects of exterior platform angle (henceforth EPA) and platform depth (henceforth PD) on flake size as observed in previous controlled experiments using glass cores (Dibble and Rezek 2009; Pelcin and Dibble 1995; Rezek et al. 2011) also hold true in similar ways for raw materials other than glass. Raw material type is therefore one of the independent variables in this experiment. Cores of four different raw materials are tested: basalt, flint, obsidian, and glass. Glass cores are molded identically to a specific size and surface morphology. The same core shapes are then produced in three additional raw materials: obsidian, basalt, and flint. Each core, however, is produced by hand, so size and shape vary minimally (Fig. 1). Cores on all four materials have an external surface with a central ridge where two faces converged at the top of the core surface. In our analysis, we exclude flakes with terminations other than feather (hinge, step, overshot) and broken flakes where the size could not be measured. Thus of 81 flakes produced in this experiment, only 57 were available for analysis (Table 1).

The raw materials come from three locations. The basalt comes from near Ashfork, Arizona, the obsidian from Jalisco, Mexico, and the flint is from Ingram, England. The raw material nodules were purchased from a flintknapping supply house (www.neolithics.com) and shaped into our specific cores there. There are microstructure differences between these three materials. Basalt is a fine-grained mafic rock, and flint is very fine cryptocrystalline quartz, while no microstructure is observed in obsidian.

Other independent variables are EPA, PD (Fig. 2), and displacement speed. All types of raw materials are cut with a

Table 1 Sample size by raw material and by exterior platform angle (EPA)

| EPA | Glass | Obsidian | Flint | Basalt |
|-----|-------|----------|-------|--------|
| 65° | 4     | 5        | 5     | 5      |
| 75° | 4     | 6        | 6     | 7      |
| 85° | 3     | 4        | 4     | 4      |
| Total | 11 | 15 | 15 | 16 | 57 |
rock saw at one end to a specific EPA. We varied EPA by three values: 65°, 75°, and 85°. During the knapping process, we held constant variables known to affect flake size and shape: angle of blow (10°), hammer tip (steel edge), and core surface morphology (no further modifications are made to the exterior surface that was pre-shaped to the center-ridge morphology shown above).

The experiment followed similar protocols to previous controlled experiments using glass alone (Dibble and Rezek 2009; Leader et al. 2017; Lin et al. 2013; Magnani et al. 2014). A core is mounted in a stable mount supported on three sides and against side rails on the exterior surface. The cores are inserted into a stable mount in an Instron Servohydraulic testing press (Model 1331). Cores are then positioned such that the angle between the platform surface and the hammer strike (angle of blow) is set to a specific value. The Instron Servohydraulic Testing Press is retrofitted with an external control panel from which displacement speed and hammer

| Site                | Level | N   | Reference                        |
|---------------------|-------|-----|----------------------------------|
| Abri Peyrony        | 1B    | 27  | (Soressi et al. 2013)            |
| Abri Peyrony        | L-3A  | 53  | (Soressi et al. 2013)            |
| Abri Peyrony        | L-3B  | 182 | (Soressi et al. 2013)            |
| Abri Peyrony        | U-2   | 35  | (Soressi et al. 2013)            |
| Abri Peyrony        | U-3A  | 140 | (Soressi et al. 2013)            |
| Abri Peyrony        | U-3C  | 6   | (Soressi et al. 2013)            |
| Combe-Capelle Bas I | I-1B  | 4   | (Dibble and Lenoir 1995)         |
| Combe-Capelle Bas I | I-1D  | 89  | (Dibble and Lenoir 1995)         |
| Combe-Capelle Bas I | I-1D1 | 26  | (Dibble and Lenoir 1995)         |
| Combe-Capelle Bas I | I-1E  | 92  | (Dibble and Lenoir, 1995)        |
| Combe-Capelle Bas I | I-2A  | 43  | (Dibble and Lenoir 1995)         |
| Combe-Capelle Bas I | I-2B  | 18  | (Dibble and Lenoir 1995)         |
| Combe-Capelle Bas II| II-4A | 28  | (Dibble and Lenoir 1995)         |
| Combe-Capelle Bas II| II-4B | 22  | (Dibble and Lenoir 1995)         |
| Combe-Capelle Bas II| II-4C | 29  | (Dibble and Lenoir 1995)         |
| Combe-Capelle Bas II| II-4D | 4   | (Dibble and Lenoir 1995)         |
| Combe-Capelle Bas II| II-4E | 24  | (Dibble and Lenoir 1995)         |
| Pech de l’Azé IV   | 3A    | 167 | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 3B    | 385 | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 4A    | 17  | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 4B    | 5   | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 4C    | 89  | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 5A    | 153 | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 5B    | 46  | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 6A    | 246 | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 6B    | 212 | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 7     | 38  | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Pech de l’Azé IV   | 8     | 148 | (Dibble et al. 2009, 2018; Turq et al. 2011) |
| Roc de Marsal       | 1     | 44  | Unpublished                      |
| Roc de Marsal       | 2     | 93  | Unpublished                      |
| Roc de Marsal       | 3     | 14  | Unpublished                      |
| Roc de Marsal       | 4     | 210 | Unpublished                      |
| Roc de Marsal       | 5     | 177 | Unpublished                      |
| Roc de Marsal       | 6     | 28  | Unpublished                      |
| Roc de Marsal       | 7     | 149 | Unpublished                      |
| Roc de Marsal       | 8     | 154 | Unpublished                      |
| Roc de Marsal       | 9     | 336 | Unpublished                      |
| Total               |       | 3533|                                  |
distance can be controlled. Displacement speed is held constant at 0.25 in./s. The hydraulics move a central actuator capable of producing 20,000 lbs of force. A steel hammer strikes the cores on the striking surface detaching the flake. Achieving a specific PD proved to be difficult. The striker rarely hits the core at the marked PD value, most likely due to the thickness of the striker, which is approximatively 1 mm. Therefore, the resulting PD shows a slight offset from the desired PD value and is thus a continuous variable that ranges from 2 to 10 mm.

Dependent variables include flake weight, volume, length, width, and thickness. The linear measurements are recorded with calipers to the nearest 0.05 mm. Platform depth is measured four times independently by four individuals and then averaged for accuracy to mitigate against inter-observer error (see Leader et al. 2017). Flakes and cores are weighed to the nearest 0.1 g. However, because the raw materials used in this experiment vary in densities, their weights are not directly comparable. Instead of weights, we use volume as a proxy for size. We first recorded the volume of a core of each raw material by measuring the volume of their 3D scans in Meshlab software. Then their densities are calculated by dividing weight by volume. These density values are then used to calculate the volumes of each flake.

To also test the relationships derived from the experimental setting against actual archeological data, a set of several archeological assemblages are used in this study. Our sample here is composed of complete flakes from three excavated Middle Paleolithic sites in southwest France: Abri Peyrony (Soressi et al. 2013), Pech de l’Azé IV (Dibble et al. 2009, 2018; Turque et al. 2011), Combe-Capelle Bas (Dibble and Lenoir 1995), and Roc de Marsal (Table 2). While some inter-observer error in the data can be expected, all the attributes are recorded in a similar manner. Included are artifacts larger than 2.5 cm in their largest dimension that have values recorded for the variables that are analyzed here. Cases with values for EPA between 50 and 90° and PD between 1.5 and 11 mm are included in the sample so as to correspond to the intervals in the experimental dataset.

Data and figures are produced in R software (R Core Team 2018). To examine the relationship between independent and dependent variables, we used linear modeling. Response variables are flake volume and linear dimensions of length, width, and thickness; and predictor variables are PD, EPA, and raw material. Volume as a response variable is transformed to its cube root so as to transform the relationship of volume, which increases in three dimensions, and PD, which increases in one dimension, to a linear one. When incorporating the interaction of EPA and PD in the model, we standardized them with the z-transformation to have a mean of 0 and standard deviation of 1. For each model, we checked for violations of its assumptions by examining the residual distribution and variance inflation factor. Influential cases were identified by Cook’s distance, leverage, and by comparing fitted values between model using all data and model with cases excluded one at a time (DFFITS). Influential cases are then are removed from the dataset. Covariates EPA and PD exhibit a high correlation in our dataset, thus violating model assumption. This correlation, however, is due to the nature of our experimental design.

| Est. coef. | Std. error | t | p | CI | β |
|-----------|------------|---|---|---|---|
| Intercept* | 17.14      | 0.35 | 48.75 | < 0.001 | 16.43–12.84 |
| PD         | 6.31       | 0.3 | 21.28 | < 0.001 | 5.72–6.91 | 1.44 |
| EPA        | 4.14       | 0.28 | 14.75 | < 0.001 | 3.58–4.71 | 0.99 |
| Flint      | 0.35       | 0.49 | 0.72 | 0.47 | −0.63–1.33 | 0.03 |
| Glass      | 0.95       | 0.52 | 1.83 | 0.07 | −0.09–2 | 0.09 |
| Obsidian   | 1.83       | 0.49 | 3.7 | 0.005 | 0.84–2.82 | 0.19 |
| PD:EPA     | 2.04       | 0.22 | 9.23 | < 0.001 | −1.6–2.48 | 0.43 |

F(6, 49) = 84.41, p < 0.001; Adj. r² = 0.9
*One influential case removed

![Fig. 3 Volume (cube root) as a function of platform depth for each exterior platform angle group, by raw material](image-url)
Because increases in both EPA and PD result in larger flakes, producing flakes with high values of both variables would produce overshot flakes. For this reason, for EPA values of 75° and 85°, PD values are in the lower range (maximum PD for 85° is 5 mm). This results in high negative correlation of EPA and PD. Therefore, we rely on variance inflation factor to detect multicollinearity of predictors. We used ANOVA to compare the null and full models when adding independent variables or interaction terms. Additional packages used in R are dplyr (Wickham et al. 2018), ggplot2 (Wickham 2016), car (Fox and Weisberg 2011), coin (Hothorn et al. 2006), and lm.beta (Behrendt 2014).

Results

Flake volume

In many previously published experiments performed on glass (Dibble and Rezek 2009; Dibble and Whittaker 1981; Pelcin and Dibble 1995; Speth 1972, 1975, 1981), flake weight was shown to be a main function of two independent variables—EPA and PD—and increasing either of these results in larger flakes. Here, we modeled the effect of raw material on the relationship of EPA-PD on flake size. We first used multiple regression to predict flake volume with EPA and PD as independent variables with the entire experimental assemblage including all raw materials. We compared the model with EPA and PD as predictors with a model that includes the interaction term of these two covariates (one influential case is excluded from the dataset). ANOVA model comparison shows that adding the interaction improves the model ($F(1,53) = 60.297, p < 0.001$). The model shows high predictability of EPA and PD on flake size ($r^2 = 0.88$, $F(3,52) = 131.9, p < 0.001$). Raw material then is added as a categorical predictor to this model and then compared to the reduced model (EPA and PD with interaction terms) with ANOVA. The results show that the raw material improves the model as well ($F(3,49) = 5.178, p = 0.003$); however,
the $r^2$ as a measure of fit shows only negligible increase ($r^2 = 0.9$, $F(6, 49) = 84.41$, $p < 0.001$). Moreover, while only obsidian shows a significant effect, confidence intervals overlapping with other raw material and low beta coefficients for obsidian (Table 3) suggest that this significance, especially given the small sample size, should be taken cautiously.

In looking at the relationship between flake volume and platform depth, cores with identical EPAs can be compared, as shown in Fig. 3. Note that each material displays the same relationship—increasing values of PD results in larger flakes. It was shown earlier (Dibble and Rezek 2009) for a separate sample of glass flakes that the relationship between flake volume and PD changes as EPA increases for all of these materials. Here (Fig. 3), slope values increase slightly for EPA of 75° and substantially for 85°.

Another way to see if flakes with different EPAs are larger or smaller relative to their PDs is to look at the ratio of the cube root of flake volume standardized by PD. Unlike the situation with EPA, it is not technically possible in our setup to produce flakes with a specific PD. For this reason, flakes are made within a wide range of PD, usually from 2 to 10 mm, depending on the value of EPA being tested (PDs that are too large will result in flakes that overshoot the core, which are not used in these analyses). As shown in Fig. 4, for all of the materials, the ratio of flake volume (cube root) to platform depth

| Table 4 Results of the regression models for predicting linear dimensions based on PD and EPA. Raw material was not a significant predictor and was excluded from the models here. Interaction term between EPA and PD for modeling thickness is omitted as it did not contribute to the model. Covariates are z-transformed to a mean of 0 and standard deviation of 1. |
|---------------------------------|-------|-----|---|---|-------|
| **Length**                     | **Est. coef.** | **Std. error** | **t** | **p**  | **CI** |
| Intercept                      | 93.3  | 2.68 | 34.8 | <0.001 | 87.88-98.63 |
| PD                             | 22.1  | 3.04 | 7.28 | <0.001 | 16.03-28.22 |
| EPA                            | 33.1  | 3.07 | 10.8 | <0.001 | 26.92-39.22 |
| PD:EPA                         | 14.5  | 2.48 | 5.86 | <0.001 | 9.55-19.52 |
| F(3, 52)=41.57, $p<0.001$; Adj. $r^2 = 0.69$ |
| **Width**                      | **Est. coef.** | **Std. error** | **t** | **p**  | **CI** |
| Intercept                      | 23.39 | 0.575 | 40.7 | <0.001 | 22.23-24.54 |
| PD                             | 9.22  | 0.705 | 13.1 | <0.001 | 7.81-10.64 |
| EPA                            | 3.2   | 0.658 | 4.86 | <0.001 | 1.87-4.51 |
| PD:EPA                         | 1.7   | 0.613 | 2.76 | <0.001 | 0.46-2.92 |
| F(3, 49)=73.3, $p<0.001$; Adj. $r^2 = 0.81$ |
| **Thickness**                  | **Est. coef.** | **Std. error** | **t** | **p**  | **CI** |
| Intercept                      | -7.53 | 2.20  | -3.42 | <0.001 | -11.95-(-3.11) |
| PD                             | 1.36  | 0.107 | 12.7 | <0.001 | 1.14-1.57 |
| EPA                            | 0.084 | 0.024 | 3.49 | <0.001 | 0.04-0.13 |
| F(2, 54)=107.7, $p<0.001$; Adj. $r^2 = 0.8$ |

Fig. 6 Dotplot showing the effect of exterior platform angle on flake length (standardized by platform depth) for different raw materials.
increases with each value of EPA. This means that at higher values of EPA, equivalent changes in PD result in proportionally larger changes in flake volume. Again, this has been repeatedly observed previously in glass alone (Dibble and Rezek 2009; Lin et al. 2013; Pelcín 1997a, b). There is less variability in the volume to platform depth at EPAs of 65° than at 75° and especially 85°. This is probably due to measurement error of PD. As EPA increases, each unit increase in PD has a more pronounced effect on flake size. Thus, small errors in measurement will result in higher standard deviations with higher EPAs.

As an additional test, we verified these results with archeological data, simply because they represent a variety of materials that were used in the past (in real-life). As is seen in Fig. 5, these assemblages demonstrate the same patterns, namely that weight is both a function of EPA and PD. Similar patterns in archeological materials were also presented in Dibble (1997).

**Individual dimensions**

Flake volume is a function of the three linear dimensions of length, width, and thickness, so it is of interest to see how these individual dimensions are themselves influenced by EPA and PD across the four material types. It has been previously reported that all three dimensions are affected by both EPA and PD, so again it is necessary to take both of these independent variables into consideration at the same time. When predicting linear dimensions with EPA and PD using multiple regression, the addition of an interaction term improved the model for length ($F(1,52) = 34.278, p < 0.001$) and width ($F(1,49) = 7.6372, p = 0.008$), but not for thickness.
Comparing those models with the models where raw material is added as a categorical covariate revealed that raw material is not a significant predictor for the linear dimensions. Figures 6, 8, and 10 show the effect of EPA on linear dimensions, standardized by PD. Similar effects are observed in archeological materials (Figs. 7, 9, 11).

Standardized coefficients ($\beta$) of the models on experimental assemblages suggest that EPA has greater effect than does PD in determining length, while the opposite is true for width and thickness—they are more affected by PD than EPA (Table 4, Figs. 6; 7; 8; 9; 10; 11). Similar trends have been observed on actualistic experimental assemblages (Dogandžić et al. 2015). With linear dimensions impacted differently by EPA and PD, the changes in the shape of the resulting flakes, e.g., elongation or relative thickness, could be predicted. In other words, particular flake characteristics can be realized by adjustments in these two platform variables (Lin et al. 2013). For instance, to achieve a long and thin blank, one has to keep the PD low while increasing the EPA. The fact that the effects of EPA and PD on linear dimensions are unsusceptible to differences in raw materials lends support for the archeological applications of different allometry ratios. In particular, the ratio of thickness, a dimension that is not affected by

**Fig. 9** Boxplots showing relationship between platform depth and width for exterior platform angle groups in archeological assemblages (Table 2)

**Fig. 10** Dotplot showing the effect of exterior platform angle on flake thickness (standardized by platform depth) for different raw materials
retouch, and length/width or weight, dimensions that reduce with reduction, can confidently be used as a measure of resharpening across different materials.

**Force**

As described in Dibble and Rezek (2009), the current experimental design, which has the steel hammer attached to a load cell, allows us to measure the load, or force, required to remove a flake. Basically, as the hammer begins to make contact with the platform surface, the load increases until the exact moment when the flake is released; at this point, the load returns to zero. In Dibble and Rezek (2009), it was shown that force is highly correlated with flake weight—larger flakes require more force for their removal—and this relationship was not affected by variation in EPA, PD, or angle of blow except for the fact that these independent variables all contribute to flake weight. For example, two flakes of the same weight can result from a high EPA and low PD or vice versa, but the force required for their removal is the same.

Force is a good predictor of weight ($r^2 = 0.38$, $p < 0.001$, $F(1,50) = 32.51$), and when raw material is added as a term in the model, it shows that it contributes to weight prediction ($r^2 = 0.55$, $p < 0.001$, $F(4,47) = 16.63$; ANOVA model comparison ($F(-3,50) = 7.26$, $p < 0.001$). This is the one comparison tested here that does show a major difference among raw materials—harder materials, flint, and basalt, require more force than do glass and obsidian (Fig. 12), as most knappers would readily attest. Our data also suggest that flint, glass, and obsidian at lower EPAs require more force to remove a flake of particular size (Fig. 12); EPA improves the model where weight is predicted by force and raw materials (ANOVA:

![Fig. 11 Boxplots showing relationship between platform depth and thickness for exterior platform angle groups in archeological assemblages (Table 2)](image)

![Fig. 12 Relationship of force required to remove flakes by weight (log transformed) for different raw materials and varying exterior platform angles](image)
relationship of force, platform variables, and size. More data are needed to improve our understanding of the experimental setting and dependent variables obtained in between independent and

The relationship

Table 5

| Independent variable            | Dependent variable                                      | Reference            |
|--------------------------------|--------------------------------------------------------|----------------------|
| EPA, PD, platform area          | Flake weight, length, width, thickness                  | (Dibble 1997)        |
| Exterior platform beveling      | Platform width, length, weight                          | (Leader et al. 2017) |
| EPA, PD                         | Flake weight, length/width, area/thickness, area/weight, usable edge, bulb length | (Lin et al. 2013)    |
| EPA, PD                         | Usable edge                                             | (Režek et al. 2018)  |
| EPA, PD, core morphology        | Flake shape                                             | (Režek et al. 2011)  |

Discussion and conclusions

The purpose of this present experiment was to address the question of whether the results of controlled experiments on glass cores are applicable to raw materials that were used in the past. We addressed this question by modeling the effect of raw materials on the relationships of EPA and PD to flake size, as measured by volume and linear dimensions. In nearly every comparison, we were unable to show that raw material has an effect on the EPA and PD relationship to size. Increasing both EPA and PD will increase flake size, regardless of the raw material type (including glass). Furthermore, the two variables interact in a way that the effect of changes in PD is amplified at higher values of EPA. Our results show that the knappers can use the same combination of EPA, PD, and core morphology to obtain similar products without making adjustments for different materials. We have shown that what makes a difference in knapping different raw materials is the amount of force needed to remove a flake. With the same core configuration, EPA and PD, the flakes struck from different raw materials will be of comparable sizes, but more force is required to remove flakes in flint and basalt than in glass or obsidian.

Some variations in flake size are noted in the experimental assemblage. For instance, the variation in standardized volume or linear dimensions is higher with higher values of EPA. Perhaps this is due to slight differences in the core surface topography—the basalt, flint, and obsidian cores were all shaped by hand by the same individual, while the glass cores were molded according to a slightly different design. Potentially, the effect of core surface morphology on size might increase with higher values of EPA. Nevertheless, the effects of EPA and PD are the same in all raw materials, and this variation appears irrespective of raw material.

It is useful to put the present study in the theoretical context of experimental design and inferential validity in archeology. In comparing the archeological record to experiments, Eren et al. 2016 argue that the former is high in external validity given that it represents a direct evidence of past behavior, but low in external validity because it is a biased and incomplete record that does not allow for control or randomization of variables. Experiments, on the other hand, have high internal validity due to their repeatability and the potential for controlling the variables, but can never precisely replicate the behavior behind the archeological record. Lin et al. 2018 use these terms in a different way when discussing the reliability of inferences derived from experiments. They state that three kinds of inferential validity play a role in experiments. The first is internal validity, which refers to the level of precision in the experimental design and the degree to which the experiment can isolate the effects of a single independent variable. The second, external validity, refers to whether the experimental results can be generalized to contexts beyond the experimental setting itself. The third kind of inferential validity is ecological validity, which basically refers to how well the results from experimental design repeat what is observed in the field. Here we use these terms as per Lin et al. 2018.

Early controlled experiments on flake production (Speth 1972; Faulkner 1972; Bonnichsen 1977; Dibble and Whittaker 1981; Pelcin and Dibble 1995, see also Dibble and Rezek 2009) tended to be reasonably high in internal validity in being able to isolate or vary in a controlled fashion the effects of any single independent variable. Unfortunately, due largely to their experimental design, they were low in ecological validity: to put it simply, the flakes produced by these experiments bore little resemblance to those found in archeological sites because the shapes of the cores used did not resemble those used in prehistoric assemblages (see Rezek et al. 2016 for review). In addition, a number of earlier experiments (Dibble and Whittaker 1981; Pelcin and Dibble 1995) were performed on plate glass, with flakes removed from the edge. While these flakes did resemble burin spalls, flake width was invariant, thus limiting the ability to generalize the results to further our understanding of how width and two-dimensional flake shape were being affected by the independent variables under study. Not surprisingly, there were few attempts to apply the findings from these experiments to archeological collections (Dibble 1981, 1997). A new experimental design, used first about 10 years ago (Dibble and Rezek 2009) and used again here, overcame some of these limitations by using glass cores that could be molded which allowed core surface configurations that yielded flakes that more closely resembled archeological ones. In itself, this new design helped to
increase both the external and ecological validity of the experiment. The use of glass cores therefore maintained high internal validity due to the fact that all of the cores were molded to be identical in size and shape. However, using glass raises a question as to the external and ecological validity of the experiments simply because glass is not a material that was used in prehistoric knapping.

Our results verified that the relationships between the platform attributes, EPA and PD on flake size, obtained from experiments on glass cores can indeed be extended to other materials that were used in the past, in this case, varieties of basalt, flint, and obsidian. This is important because it shows that experiments using glass cores can be generalized and extended to other materials. In other words, they also have a relatively high degree of external validity. We note that, besides glass, porcelain is also gaining in popularity for knapping experiments (e.g., Khreisheh et al. 2013), and this material too should be evaluated to ensure its external and ecological validity and therefore suitability as a test material.

There are similarities in how raw materials tested in these experiments behave during knapping and potentially these results can be generalized to all knappable materials. There is a possibility, however, that other—or even many—materials used in the past may respond differently. Moreover, not all flints, basalts, and obsidians are the same, and so within these broad categories of rocks, considerable variability exists that may also affect their fracture properties. While it is impossible to perform experiments on every variety of rock that exists in nature, a more fruitful approach would be to study fracture patterns in relation to basic properties of materials—such as hardness, brittleness, and homogeneity—which could ultimately enable us to generalize experimental results to a much larger degree and model how those properties might affect particular properties of lithic assemblages. Such an approach is beyond the scope of the present paper, but it does represent a fruitful avenue for future research. We will reiterate, however, that the three raw materials tested here show a degree of variation in their properties and yet respond to changes in EPA and PD in a comparable way, and in a comparable way also to glass.

Comparisons with archeological materials allow us to assess the ecological validity of these experimental results. The general finding is that the effects of EPA and PD are observable in archeological assemblages as well. However, note that, again, even in these samples there was a limited variety of raw materials—they were mostly Late Cretaceous flints. While the comparisons presented here are concerned only with overall size and the three linear flake dimensions, previous published experiments have presented other such comparisons with reference to archeological data (see Table 5).

The applicability of the results from the experiments presented here to archeological context is reflected in the corroboration of the methods for estimating the original size of already reduced and discarded artifacts found in the archeological record (Braun et al. 2008; Dibble 1997; Lin et al. 2013). In that regard, the convergence of results stemming from controlled and free-hand (Clarkson and Hiscock 2011; Davis and Shea 1998; Doganžić et al. 2015; Muller and Clarkson 2014, 2016; Shott et al. 2000) experiments on the relationship of platform variables and size speaks further to the level of ecological validity.

The clear conclusion—both from our experiments and from comparison with archeological data—is that the EPA and PD predict flake size irrespective of the raw material—including glass. Our results show that the experiments performed on glass are valid and therefore generalizable not only to several other materials that were knapped in the past but that this also provides insights into significant aspects of prehistoric lithic assemblage variability (see also Lin et al. 2013; Režek et al. 2018). We showed that it takes different amounts of force to remove flakes of the same size (weight) made on materials of different properties. We did not, however, investigate how other aspects of flaking, such as hammer type, and angle of blow, might affect the results, across different raw materials. There is undoubtedly still much to be learned about the fracture properties of various materials, but there is little doubt that a better understanding of how materials break will significantly aid in developing accurate interpretations of prehistoric chipped stone industries.

Acknowledgments The project was initially conceived of by Harold Dibble. He wrote an earlier version of this manuscript but passed away before commenting on the final manuscript. His lab misses him greatly. The petrographic analysis was done at the Center for the Analysis of Archeological Materials at the Penn Museum. We thank Craig Ratzat from www.neolithic.com for providing us with raw materials and Alain Tuq for information on the flint we used. We thank two anonymous reviewers for constructive comments on this manuscript.

Funding information Open access funding provided by Projekt DEAL. The raw material experiments were funded by the National Science Foundation (BCS 0649763 and BCS 1153192), the University of Pennsylvania Museum of Archeology and Anthropology, and the Institutional Strategy of the University of Tübingen (Deutsche Forschungsgemeinschaft, ZUK 63). T. Doganžić received funding from the European Union’s Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Skłodowska-Curie grant agreement No. 751125 for the project MORPHOLITHEX. C. Tennie was also funded by the STONECULT project. The project STONECULT has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme grant agreement No. 714658.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
Magnani M, Rezek Z, Lin SC, Chan A, Dibble HL (2014) Flake variation in relation to the application of force. J Archaeol Sci 46:37–49. https://doi.org/10.1016/j.jas.2014.02.029

Manninen MA (2016) The effect of raw material properties on flake and flake-tool dimensions: a comparison between quartz and chert. Quat Int 424:24–31. https://doi.org/10.1016/j.quaint.2015.12.096

Mercieca A (2000) Burnt and broken: an experimental study of heat fracturing in silcrete. Aust Archaeol 51:40–47. https://doi.org/10.1080/01977261.1999.11681679

Muller A, Clarkson C (2014) Estimating original flake mass on blades using 3D platform area: problems and prospects. J Archaeol Sci 52:31–38. https://doi.org/10.1016/j.jas.2014.08.025

Muller A, Clarkson C (2016) A new method for accurately and precisely measuring flake platform area. J Archaeol Sci Rep 8:178–186. https://doi.org/10.1016/j.jasrep.2016.06.015

Pelcin AW (1997a) The effect of core surface morphology on flake attributes: evidence from a controlled experiment. J Archaeol Sci 24:749–756. https://doi.org/10.1006/jasc.1996.0156

Pelcin AW (1997b) The formation of flakes: the role of platform thickness and exterior platform angle in the production of flake initiations and terminations. J Archaeol Sci 24:1107–1113. https://doi.org/10.1006/jasc.1996.0190

Pelcin AW (1998) The threshold effect of platform width: a reply to Davis and Shea. J Archaeol Sci 25:615–620. https://doi.org/10.1006/jasc.1997.0253

Pelcin A, Dibble HL (1995) The effect of hammer mass and velocity on flake mass. J Archaeol Sci 22:429–439. https://doi.org/10.1006/jasc.1995.0042

R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rezek Z, Lin S, Iovita R, Dibble HL (2011) The relative effects of core surface morphology on flake shape and other attributes. J Archaeol Sci 38:1346–1359. https://doi.org/10.1016/j.jas.2011.01.014

Rezek Z, Lin S, Dibble HL (2016) The role of controlled experiments in understanding variation in flake production. In: Sullivan AP, Olszewski DJ (eds) Archaeological variability and interpretation in global perspective. University Press of Colorado, pp 307–320. https://doi.org/10.5876/9781607324942.014

Režek Ž, Dibble HL, McPherron SP, Braun DR, Lin SC (2018) Two million years of flaking stone and the evolutionary efficiency of stone tool technology. Nat Ecol Evol 2:628–633. https://doi.org/10.1038/s41559-018-0488-4

Rick JW, Chappell S (1983) Thermal alteration of silica materials in technological and functional perspective. Lithic Technol 12:69–70. https://doi.org/10.1080/01977261.1983.11760621

Roth BJ, Dibble HL (1998) Production and transport of blanks and tools at the French middle Paleolithic site of Combe-Capelle Bas. Am Antiq 63:47–62

Schmidt P, Slodczyk A, Léa V, Davidson A, Puuad S, Scian P (2013) A comparative study of the thermal behaviour of length-fast chalcedony, length-slow chalcedony (quartzine) and moganite. Phys Chem Miner 40:331–340. https://doi.org/10.1007/s00269-013-0574-8

Shott MJ, Bradbury AP, Carr PJ, Odell GH (2000) Flake size from platform attributes: predictive and empirical approaches. J Archaeol Sci 27:877–894. https://doi.org/10.1006/jasc.1999.0499

Soressi M, McPherron SP, Lenoir M, Dogandžić T, Goldberg P, Jacobs Z, Maigrot Y, Martisius NL, Miller CE, Rendu W, Richards M, Skinner MM, Steele TE, Talamo S, Texier J-P (2013) Neandertals made the first specialized bone tools in Europe. Proc Natl Acad Sci U S A 110:14186–14190. https://doi.org/10.1073/pnas.1302730110

Speith JD (1972) Mechanical basis of percussion flaking. Am Antiq 37:34–60. https://doi.org/10.2307/278884

Speith JD (1975) Miscellaneous studies in hard-hammer percussion flaking: the effects of oblique impact. Am Antiq 40:203–207

Speith JD (1981) The role of platform angle and core size in hard-hammer percussion flaking. Lithic Technol 10:16–21. https://doi.org/10.1080/01977261.1981.11720840

Turg A (1989) Approche technologique et économique du faciès Moustérien de type Quina : étude préliminaire. Bull Société Préhistorique Fr 86:244–256. https://doi.org/10.3406/bspf.1989.9390

Turg A, Dibble HL, Goldberg P, McPherron SP, Sandgate D, Jones H, Maddison K, Maureille B, Menzer S, Rink J, Steenhuyse A (2011) Les Fouilles Récentes du Pech de l’Azé IV (Dordogne). Gall Préhistoire 53:1–58

Webb JA, Domanski M (2008) The relationship between lithology, flaking properties and artifact manufacture for Australian Silcretes. Archaeometry 50:555–575. https://doi.org/10.1111/j.1475-4754.2007.00381.x

Whittaker JC (1994) Flintknapping: making and understanding stone tools. University of Texas Press, Austin

Wickham H, François R, Henry L, Müller K (2018) dplyr: elegant graphics for data analysis. Springer-Verlag, New York

Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York

Wickham H, François R, Henry L, Müller K (2018) dplyr: a grammar of data manipulation. R package version 0.7.6

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.