## Supplemental Material

CBE-Life Sciences Education
Connell et al.

Supplementary Information for the Confirmatory Factor Analysis.

## CFA Methods

Before conducting the CFA, we analyzed our data to determine whether they were appropriate for CFA following the guidelines of Knekta et al. (2019). We first checked the data for missing values and determined whether missingness was random using Little's missing completely at random (MCAR) test, implemented with the naniar package in R (Tierney and Cook 2023). We then inspected the SAGE questions for univariate outliers by looking at mean, maximum, and minimum values for each item in the SAGE. Multivariate outliers were assessed using Mahalonobis distance. To assess factorability, we first used an inter-item correlation matrix to ensure there were sizable correlations between items, especially within a factor. We then tested for factorability using Kaiser's measure of sampling adequacy with the psych package in R (Revelle 2017). We next assessed items for univariate normality by assessing skewness and kurtosis for each item, and we used Mardia's multivariate normality test to test for multivariate normality, implemented with the psych package (Revelle 2017). Finally, we investigated multicollinearity by examining inter-item correlations and from tolerance values from multiple regressions using the olsrr package in R (Hebbali 2018).

To conduct the CFA and to determine whether students in different treatments initially responded to SAGE questions in a similar manner, we assessed measurement invariance between students in the two treatments (permanent groups and nonpermanent groups). We first fit a configural invariance model (a multigroup CFA model) to determine whether the same factors hold across groups. We then fit a metric invariance model and compared it to the configural model. Finally, we fit a scalar invariance model and compared it to the metric model. We used the lavaan package in R (Rosseel 2012) to conduct the CFA and the semTools package (Jorgenson et al. 2022) to conduct the model comparisons. We employed robust maximum-likelihood estimation (MLR) methods to account for multivariate non-normality and non-linearity (see Results). We also used full-information maximum likelihood (FIML) to handle missing data. To evaluate model fit, we used multiple fit indices (chi-square from robust MLR; comparative fit index [CFI]; root-mean-square error of approximation [RMSE]; and the standardized root-mean-square residual [SRMR]). We evaluated these indices against the following criteria: CFI $>0.95$, RMSEA $<0.06$, and SRMR < 0.08 (Hu and Bentler, 1999; Knekta et al. 2019).

## CFA Results

The preSAGE data from the group permanence study were appropriate for CFA. No items had more than $0.3 \%$ missing data ( 2 of 619 missing answers) and no student had more than three missing answers. The data were missing completely at random for the preSAGE data $\left(X^{2}=596, \mathrm{df}=651, \mathrm{p}=\right.$ $0.94)$. Regarding univariate outliers, means for the SAGE items ranged from $2.25-4.53$. Students with high Mahalanobis distance ( $\mathrm{p}<0.001$ ) were inspected. One student was removed from the analysis for obvious string responding (all responses were the same two numbers at the edge of the Likert scale, including items that were reverse coded). Inter-item correlations had several correlations >0.3 for items within a factor and Kaiser's measure of sampling adequacy was 0.91 (this value should be $>0.6$ ), indicating the data had good factorability. All items had skewness < |1.3| and kurtosis was < $|2|$ for most items and $<|4|$ for all items. This indicates most items were univariate normal, although some showed slight non-normality. However, Mardia's multivariate normality test indicated the data had significant multivariate skewness and kurtosis, so we used robust estimation methods in the CFA. The data did not show evidence of multicollinearity, as the highest correlation in the inter-item correlation matrices was 0.66 and the lowest tolerance value from the multiple regressions was 0.39 (this value should be >0.1).

The confirmatory factor analyses of the SAGE data indicated that the four SAGE constructs were modeled reasonably well in our student population and that the constructs were modeled similarly between students in permanent and nonpermanent groups. The chi-square test of model fit was significant ( $\mathrm{X}^{2}=2399, \mathrm{df}=854, \mathrm{p}<001$ ), but this test known to be very sensitive to sample size and as such is likely not a good measure of model fit in our analyses. Of the other three indices, two indicated good model fit
$($ RMSEA $=0.059$ and $\operatorname{SRMR}=0.075)$, although the CFI was lower than recommended $(\mathrm{CFI}=0.77)$. The data exhibited metric invariance (difference in chi-square between configural and metric models $=38.0$, difference in degrees of freedom $=39, p=0.52$ ) and scalar invariance (difference in chi-square between metric and scalar models $=34.7$, difference in degrees of freedom $=39, p=0.67$ ).

Factor loading of each SAGE question and correlations between factors for CFAs conducted on SAGE data from the two treatment groups (permanent and nonpermanent groups) can be found below. The number associated with each item represents the question number on the SAGE. For example, "Qual1" is SAGE question \#1 which was associated with the Quality of Product factor. The covariances between factors from the CFAs for the two treatment groups are also presented below.

## Permanent Groups

Latent Variables:

|  | Estimate | Std.Err |
| :---: | :---: | :---: |
| Quality of Product |  |  |
| Qual1 | 1.000 |  |
| Qual5 | 0.881 | 0.091 |
| Qual7 | 1.091 | 0.092 |
| Qual8 | 0.618 | 0.084 |
| Qual12 | 0.982 | 0.077 |
| Qual13 | 1.003 | 0.087 |
| Qual14 | 0.659 | 0.083 |
| Qual16 | 0.678 | 0.109 |
| Qual30 | 1.101 | 0.088 |
| Qual31 | 0.506 | 0.079 |
| Qual37 | 1.074 | 0.081 |
| Qual40 | 1.047 | 0.097 |
| Qual41 | 1.031 | 0.087 |
| Qual47 | 1.109 | 0.087 |
| Qual48 | 0.866 | 0.105 |
| Frustration |  |  |
| Frus4 | 1.000 |  |
| Frus27 | 1.086 | 0.201 |
| Frus33 | 1.006 | 0.210 |
| Frus43 | 0.823 | 0.187 |
| Frus46 | 1.384 | 0.235 |
| Frus50 | 1.096 | 0.236 |
| Frus53 | 1.126 | 0.242 |
| Frus54 | 1.027 | 0.146 |
| Peer Support |  |  |
| Peer6 | 1.000 |  |
| Peer10 | 1.092 | 0.195 |
| Peer11 | 1.141 | 0.178 |
| Peer17 | 1.137 | 0.169 |
| Peer20 | 1.463 | 0.251 |
| Peer26 | 0.947 | 0.111 |
| Peer32 | 0.838 | 0.109 |
| Peer34 | 0.908 | 0.150 |
| Interdependence |  |  |
| Int9 | 1.000 |  |
|  |  |  |


| Int19 | 0.970 | 0.269 |
| :--- | :--- | :--- |
| Int23 | 1.004 | 0.323 |
| Int25 | 0.781 | 0.195 |
| Int28 | 0.508 | 0.194 |
| Int29 | 1.350 | 0.203 |
| Int36 | 1.208 | 0.399 |
| Int38 | 0.609 | 0.155 |
| Int44 | 0.980 | 0.288 |
| Int45 | 0.473 | 0.176 |
| Int49 | 0.735 | 0.191 |
| Int52 | 1.322 | 0.396 |

Covariances:

|  | Estimate | Std.Err |
| :--- | :--- | :--- |
| Quality of Product |  |  |
| Frustration | 0.119 | 0.027 |
| Peer Support | 0.147 | 0.028 |
| $\quad$ Interdependence | 0.146 | 0.039 |
| Frustration | 0.088 | 0.023 |
| Peer Support | 0.057 | 0.021 |
| Interdependence <br> Peer Support <br> Interdependence | 0.084 | 0.026 |

## Nonpermanent Groups

Latent Variables:

|  | Estimate | Std.Err |
| :--- | :--- | :--- |
| Quality of Product |  |  |
| Qual1 | 1.000 |  |
| Qual5 | 0.807 | 0.112 |
| Qual1 | 1.068 | 0.097 |
| Qual8 | 0.442 | 0.076 |
| Qual12 | 0.811 | 0.093 |
| Qual13 | 0.910 | 0.104 |
| Qual14 | 0.635 | 0.089 |
| Qual16 | 0.600 | 0.103 |
| Qual30 | 1.016 | 0.100 |
| Qual31 | 0.576 | 0.103 |
| Qual37 | 1.107 | 0.083 |
| Qual40 | 1.173 | 0.093 |
| Qual41 | 1.139 | 0.095 |
| Qual47 | 1.139 | 0.091 |
| Qual48 | 1.042 | 0.116 |
| Frustration |  |  |
| Frus4 | 1.000 |  |
| Frus27 | 2.128 | 1.183 |
| Frus33 | 2.828 | 1.335 |
| Frus43 | 2.259 | 0.974 |
| Frus46 | 3.412 | 1.706 |


| Frus50 | 2.787 | 1.248 |
| :---: | :---: | :---: |
| Frus53 | 2.601 | 1.240 |
| Frus54 | 1.490 | 0.501 |
| Peer Support |  |  |
| Peer6 | 1.000 |  |
| Peer10 | 0.951 | 0.261 |
| Peer11 | 1.326 | 0.250 |
| Peer17 | 1.426 | 0.304 |
| Peer20 | 1.698 | 0.297 |
| Peer26 | 1.332 | 0.204 |
| Peer32 | 1.161 | 0.224 |
| Peer34 | 1.025 | 0.230 |
| Interdependence |  |  |
| Int9 | 1.000 |  |
| Int19 | 0.868 | 0.180 |
| Int23 | 1.021 | 0.241 |
| Int25 | 0.489 | 0.133 |
| Int28 | 0.391 | 0.173 |
| Int29 | 1.118 | 0.117 |
| Int36 | 1.177 | 0.293 |
| Int38 | 0.318 | 0.134 |
| Int44 | 0.858 | 0.223 |
| Int45 | 0.403 | 0.154 |
| Int49 | 0.447 | 0.144 |
| Int52 | 1.293 | 0.295 |

Covariances:

|  | Estimate | Std.Err |
| :--- | :--- | :--- |
| Quality of Product | 0.032 | 0.017 |
| Frustration | 0.080 | 0.018 |
| Peer Support | 0.152 | 0.040 |
| $\quad$ Interdependence |  |  |
| Frustration 0.021 <br> $\quad$ Peer Support 0.009 <br> $\quad$ Interdependence 0.026 <br> Peer Support <br> Interdependence 0.059$\quad 0.018$ |  |  |

## References

Hebbali, A. (2018) olsrr: Tools for building OLS Regression Models, version 0.5.1. Retrieved from https://cran.r-project.org/web/packages/olsrr.
Hu, L. \& Bentler, P. (a999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternative. Structural Equation Modeling: A Multidisciplinary Journal 6, 1-55
Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., \& Rosseel, Y. (2022). semTools: Useful tools for structural equation modeling. R package version 0.5-6. Retrieved from https://CRAN.Rproject.org/package=semTools

Revelle, W. (2017). psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA. Retrieved from https://CRAN.Rproject.org/package=psychVersion=1.7.8
Rosseel, Y. (2012) lavaan: An R package for structural equation modeling. Journal of Statistical Software 48, 1-36.
Tierney, N. \& Cook, D. (2023) Expanding tidy data principles to facilitate missing data exploration, visualization and assessment of imputations. Journal of Statistical Software. doi: 10.18637/jss.v105.i07

Supplementary Table 1. Codes derived from responses to the open-ended question: "Do you prefer to stay in the same group for the entire quarter or change groups? Please explain." Examples of student responses are provided. Two coders analyzed the responses after co-coding 70 responses from each group type (permanent and nonpermanent) and reaching consensus on the codes. The final coding included 160 overlapping responses ( 80 from each group type), which were used to calculate Cohen's kappa for each code.
Code
Liked/disliked group
"Change because I wouldn't have wanted to be stuck with the same
people if my group was not a good group."
$\frac{\text { Cohen's kappa }}{0.72}$
"Change because I wouldn't have wanted to be stuck with the same people if my group was not a good group."
"I loved my group. I would've been sad to switch groups."
Positive group interactions
"I prefer to stay in the same group. You get to know each other better and it gives us the opportunity to become more comfortable asking questions. It is encouraging to have the same people learn and grow alongside you. It also helps us to know who has what strengths on group tests."

Startup costs/logistics
"Same group, it would be challenging to switch groups since being more of an introvert it is hard to get comfortable with people."
"Same group so we don't have to waste time on formalities every so often."
Group disfunction
"I think that last module was really difficult due to the fact that my group had become friends and wanted to use the time to talk instead of listen to the lecture. I think it would have been more beneficial change groups."
"I would prefer to change groups because we didn't work together."
Different ideas \& perspectives
"I liked mixing up the groups. I felt like I got to work with a lot of people with varying biology knowledge and it was helpful in the long run to gain so many new perspectives that I may not have thought of myself."
"I liked to change groups. It was a good way to meet new people and we got to hear a variety of opinions."
Change is good/bad
"Changing groups helps keep things fresh."
"Same group, change ain't great when it comes to group work."

Supplementary Table 2. Best-fit models for the content assessment and the four SAGE factors, including demographic factors, when students were in small and large groups.

|  | Best-fit model | Estimate $\pm$ SE | $t$ or $z$ value* |
| :---: | :---: | :---: | :---: |
| Individual Content Assessment | post $\sim$ pre + GPA |  |  |
|  | intercept | $4.85 \pm 1.16$ | 4.20 |
|  | preassessment | $0.58 \pm 0.06$ | 10.21 |
|  | GPA | $3.41 \pm 0.32$ | 10.53 |
| Gender | post $\sim$ pre + gender + GPA |  |  |
|  | intercept | $4.51 \pm 1.16$ | 4.43 |
|  | preassessment | $0.56 \pm 0.06$ | 9.91 |
|  | gender (ref: female) | $0.93 \pm 0.43$ | 2.15 |
|  | GPA | $3.51 \pm 0.32$ | 10.89 |
| BIPOC status | post $\sim$ pre + BIPOC $+G P A$ |  |  |
|  | intercept | $5.12 \pm 1.16$ | 3.89 |
|  | preassessment | $0.58 \pm 0.06$ | 10.34 |
|  | BIPOC status (ref: not BIPOC) | $-1.11 \pm 0.50$ | 2.22 |
|  | GPA | $3.38 \pm 0.32$ | 10.49 |
| First generation status | post $\sim$ pre $+G P A$ |  |  |
|  | intercept | $4.85 \pm 1.16$ | 4.20 |
|  | preassessment | $0.58 \pm 0.06$ | 10.21 |
|  | GPA | $3.41 \pm 0.32$ | 10.53 |
| SAGE Quality of Product | postQual ~ preQual |  |  |
|  | intercept | $17.09 \pm 2.42$ | 7.05 |
|  | preQuality | $0.72 \pm 0.05$ | 15.62 |
| Gender | postQual ~ preQual |  |  |
|  | intercept | $17.09 \pm 2.42$ | 7.05 |
|  | preQuality | $0.72 \pm 0.05$ | 15.62 |
| BIPOC status | postQual ~ preQual |  |  |
|  | intercept | $17.09 \pm 2.42$ | 7.05 |
|  | preQuality | $0.72 \pm 0.05$ | 15.62 |
| First generation status | postQual ~ preQual |  |  |
|  | intercept | $17.09 \pm 2.42$ | 7.05 |
|  | preQuality | $0.72 \pm 0.05$ | 15.62 |
| SAGE Peer Support | $\begin{aligned} & \text { postPeer } \sim \text { prePeer }+G P A+1 \mid \text { group } \\ & \text { prePeer Support } \end{aligned}$ | $0.24 \pm 0.03$ | 7.85 |
|  | GPA | $0.45 \pm 0.17$ | 2.73 |
| Gender | $\begin{aligned} & \text { postPeer } \sim \text { prePeer }+G P A+1 \mid \text { group } \\ & \text { prePeer Support } \end{aligned}$ | $0.24 \pm 0.03$ | 7.85 |
|  | GPA | $0.45 \pm 0.17$ | 2.73 |
| BIPOC status | $\begin{aligned} & \text { postPeer } \sim \text { prePeer }+G P A+1 \mid \text { group } \\ & \text { prePeer Support } \end{aligned}$ | $0.24 \pm 0.03$ | 7.85 |
|  | GPA | $0.45 \pm 0.17$ | 2.73 |


| First generation status | postPeer $\sim$ prePeer $+G P A+1 \mid$ group prePeer Support GPA | $\begin{aligned} & 0.24 \pm 0.03 \\ & 0.45 \pm 0.17 \end{aligned}$ | $\begin{aligned} & 7.85 \\ & 2.73 \end{aligned}$ |
| :---: | :---: | :---: | :---: |
| SAGE Interdependence | postInt~preInt $+1 \mid$ group preInterdependence | $0.26 \pm 0.03$ | 9.93 |
| Gender | postInt $\sim$ preInt $+1 \mid$ group preInterdependence | $0.26 \pm 0.03$ | 9.93 |
| BIPOC status | postInt $\sim$ preInt $+1 \mid$ group preInterdependence | $0.26 \pm 0.03$ | 9.93 |
| First generation status | postInt $\sim$ preInt $+1 \mid$ group preInterdependence | $0.26 \pm 0.03$ | 9.93 |
| SAGE Satisfaction with Group | postSat $\sim$ preSat $+1 \mid$ group preSatisfaction | $0.26 \pm 0.03$ | 9.12 |
| Gender | postSat $\sim$ preSat $+1 \mid$ group preSatisfaction | $0.26 \pm 0.03$ | 9.12 |
| BIPOC status | postSat $\sim$ preSat $+1 \mid$ group preSatisfaction | $0.26 \pm 0.03$ | 9.12 |
| First generation status | postSat $\sim$ preSat $+1 \mid$ group preSatisfaction | $0.26 \pm 0.03$ | 9.12 |

*The content assessment models were estimated using the lmer function, with the lme4 package in R (Bates, Maechler, Bolker, \& Walker 2015), with returns a $t$ value. Models of SAGE constructs were estimated using the clmm function, with the ordinal package in R (Christensen 2018) to account for the Likert scale data, which returns a $z$ value. The clmm function does not return a model intercept, so those have not been reported when clmm required for the best-fit model. The critical value for t -values and z values is identical; values of 1.96 are considered "statistically significant" to $\mathrm{p}<0.05$ but note that interpreting p -values after model selection is performed is not advised.

Supplementary Table 3. Best-fit models for the content assessment and the four SAGE constructs, including demographic factors, when students were in permanent and nonpermanent groups.

| Content assessment | Best-fit model <br> post $\sim$ pre $+G P A+(1 \mid$ section $)$ | Estimate $\pm$ SE | $t$ or $z$ value* |
| :---: | :---: | :---: | :---: |
|  | intercept | $5.54 \pm 1.05$ | 5.27 |
|  | preassessment | $0.54 \pm 0.05$ | 12.10 |
|  | GPA | $3.58 \pm 0.29$ | 12.50 |
| Gender | post $\sim$ pre + gender $+G P A+(1 \mid$ section $)$ |  |  |
|  | intercept | $5.14 \pm 1.05$ | 4.88 |
|  | preassessment | $0.52 \pm 0.05$ | 11.42 |
|  | gender (ref: female) | $1.07 \pm 0.35$ | 3.06 |
|  | GPA | $3.70 \pm 0.29$ | 12.92 |
| BIPOC status | post $\sim$ pre $+G P A+(1 \mid$ section $)$ |  |  |
|  | intercept | $5.54 \pm 1.05$ | 5.27 |
|  | preassessment | $0.54 \pm 0.05$ | 12.10 |
|  | GPA | $3.58 \pm 0.29$ | 12.50 |
| First generation status | post $\sim$ pre + first.gen $+G P A+(1 \mid$ section $)$ |  |  |
|  | intercept | $6.38 \pm 1.09$ | 5.84 |
|  | preassessment | $0.53 \pm 0.05$ | 11.68 |
|  | first gen status (ref: not first gen) | $-0.93 \pm 0.38$ | 2.48 |
|  | GPA | $3.47 \pm 0.29$ | 12.01 |
| SAGE Quality of Product | $\begin{aligned} & \text { postQual } \sim \text { preQual }+(1 \mid \text { group }) \\ & \text { preQuality } \end{aligned}$ | $0.18 \pm 0.01$ | 15.67 |
| Gender | $\begin{aligned} & \text { postQual } \sim \text { preQual }+(1 \mid \text { group }) \\ & \text { preQuality } \end{aligned}$ | $0.18 \pm 0.01$ | 15.67 |
| BIPOC status | $\begin{aligned} & \text { postQual } \sim \text { preQual }+ \text { BIPOC }+(1 \mid \text { group }) \\ & \text { preQuality } \end{aligned}$ | $0.18 \pm 0.01$ | $15.73$ |
|  |  | $-0.38 \pm 0.18$ | $2.15$ |
| First generation status | $\begin{aligned} & \text { postQual } \sim \text { preQual }+(1 \mid \text { group }) \\ & \text { preQuality } \end{aligned}$ | $0.18 \pm 0.01$ | 15.67 |
| SAGE Peer Support | $\begin{aligned} & \text { postPeer } \sim \text { prePeer }+ \text { permanence }+(1 \mid \text { group }) \\ & \text { prePeer } \end{aligned}$ | $0.23 \pm 0.02$ | 9.81 |
|  | permanence (ref: nonpermanent) | $0.44 \pm 0.17$ | 2.68 |
| Gender | $\begin{aligned} & \text { postPeer } \sim \text { prePeer }+ \text { permanence }+(1 \mid \text { group }) \\ & \text { prePeer } \end{aligned}$ | $0.23 \pm 0.02$ | 9.81 |
|  |  | $0.44 \pm 0.17$ | 2.68 |
| BIPOC status | $\begin{aligned} & \text { postPeer } \sim \text { prePeer }+ \text { permanence }+ \text { BIPOC }+ \\ & (1 \mid \text { group }) \end{aligned}$ |  |  |
|  | prePeer | $0.24 \pm 0.02$ | 9.97 |
|  | permanence (ref: nonpermanent) | $0.43 \pm 0.17$ | 2.61 |
|  | BIPOC status (ref: not BIPOC) | $-0.50 \pm 0.18$ | 2.78 |
| First generation status | $\begin{aligned} & \text { postPeer } \sim \text { prePeer }+ \text { permanence }+(1 \mid \text { group }) \\ & \text { prePeer } \end{aligned}$ | $0.23 \pm 0.02$ | 9.81 |


|  | permanence (ref: nonpermanent) | $0.44 \pm 0.17$ | 2.68 |
| :---: | :---: | :---: | :---: |
| SAGE Interdependence | postInt $\sim$ preInt + permanence $+(1 \mid$ group $)$ preInterdependence permanence (ref: nonpermanent) | $\begin{aligned} & 0.29 \pm 0.02 \\ & 0.66 \pm 0.18 \end{aligned}$ | $\begin{gathered} 14.47 \\ 3.64 \end{gathered}$ |
| Gender | $\begin{gathered} \text { postInt } \sim \text { preInt }+ \text { permanence }+(1 \mid \text { group }) \\ \text { preInterdependence } \\ \text { permanence }(\text { ref: nonpermanent }) \end{gathered}$ | $\begin{aligned} & 0.29 \pm 0.02 \\ & 0.66 \pm 0.18 \end{aligned}$ | $\begin{gathered} 14.47 \\ 3.64 \end{gathered}$ |
| BIPOC status | $\begin{aligned} & \text { postInt } \sim \text { preInt }+ \text { permanence }+(1 \mid \text { group }) \\ & \text { preInterdependence } \\ & \text { permanence }(\text { ref: nonpermanent }) \end{aligned}$ | $\begin{aligned} & 0.29 \pm 0.02 \\ & 0.66 \pm 0.18 \end{aligned}$ | $\begin{gathered} 14.47 \\ 3.64 \end{gathered}$ |
| First generation status | postInt $\sim$ preInt + permanence $+(1 \mid$ group $)$ preInterdependence permanence (ref: nonpermanent) | $\begin{aligned} & 0.29 \pm 0.02 \\ & 0.66 \pm 0.18 \end{aligned}$ | $\begin{gathered} 14.47 \\ 3.64 \end{gathered}$ |
| SAGE Satisfaction with group | $\begin{aligned} & \text { postSat } \sim \text { preSat }+ \text { permanence }+G P A \\ & \text { intercept } \\ & \text { preSatisfaction } \\ & \text { permanence (ref: nonpermanent) } \\ & \text { GPA } \end{aligned}$ | $\begin{gathered} 12.65 \pm 1.16 \\ 0.59 \pm 0.04 \\ 2.23 \pm 0.29 \\ -0.51 \pm 0.22 \end{gathered}$ | $\begin{aligned} & 10.91 \\ & 15.57 \\ & 7.55 \\ & 2.37 \end{aligned}$ |
| Gender | $\begin{aligned} & \text { postSat } \sim \text { preSat }+ \text { permanence }+G P A \\ & \quad \text { intercept } \\ & \text { preSatisfaction } \\ & \text { permanence (ref: nonpermanent) } \\ & \text { GPA } \end{aligned}$ | $\begin{gathered} 12.65 \pm 1.16 \\ 0.59 \pm 0.04 \\ 2.23 \pm 0.29 \\ -0.51 \pm 0.22 \end{gathered}$ | $\begin{aligned} & 10.91 \\ & 15.57 \\ & 7.55 \\ & 2.37 \end{aligned}$ |
| BIPOCstatus | postSat $\sim$ preSat + permanence + BIPOC + GPA <br> intercept <br> preSatisfaction <br> permanence (ref: nonpermanent) <br> BIPOC status (ref: not BIPOC) <br> GPA | $\begin{gathered} 12.94 \pm 1.16 \\ 0.59 \pm 0.04 \\ 2.21 \pm 0.29 \\ -0.82 \pm 0.35 \\ -0.57 \pm 0.22 \end{gathered}$ | $\begin{gathered} 11.14 \\ 15.71 \\ 7.54 \\ 2.37 \\ 2.65 \end{gathered}$ |
| First generation status | postSat $\sim$ preSat + permanence + GPA <br> intercept <br> preSatisfaction <br> permanence (ref: nonpermanent) <br> GPA | $\begin{gathered} 12.65 \pm 1.16 \\ 0.59 \pm 0.04 \\ 2.23 \pm 0.29 \\ -0.51 \pm 0.22 \end{gathered}$ | $\begin{aligned} & 10.91 \\ & 15.57 \\ & 7.55 \\ & 2.37 \end{aligned}$ |

*The content assessment models were estimated using the lmer function, with the lme4 package in R (Bates, Maechler, Bolker, \& Walker 2015), with returns a $t$ value. Models of SAGE constructs were estimated using the clmm function, with the ordinal package in R (Christensen 2018) to account for the Likert scale data, which returns a $z$ value. The clmm function does not return a model intercept, so those have not been reported when clmm required for the best-fit model. The critical value for t -values and z values is identical; values of 1.96 are considered "statistically significant" to $\mathrm{p}<0.05$ but note that interpreting p -values after model selection is performed is not advised.

Pre-Assessment, 5 extra credit pts. This assessment is scored based on completion, not accuracy of answers and will be used to gauge where the class is on various biological topics. Please answer each question to the best of your ability. Make a best guess when you do not know the answer. You may write on this assessment form.

Please turn in both this assessment form and your scantron once you are finished.

## Membrane Structure and Function.

1. All cell membranes $\qquad$ .
a. allow free movement of materials into or out of the cell.
b. allow some substances to enter the cell but prevent all substances from leaving.
c. allow only beneficial materials to enter the cell.
d. allow some substances to pass through, but not others.
2. A phospholipid molecule is diagrammed at the right, and the four diagrams A-D below represent cross sections of spherical structures composed of phospholipids. Which of these structures is most likely to form when phospholipids are vigorously dispersed in water?

3. Which is a correct diagram of how water molecules bond with one another?
d.

a.

b.

c.

4. Which is a correct diagram of how a water molecule would interact with a phospholipid? The phospholipid head
has a negative charge.

b.

c.

d.

5. Which of these substances requires a membrane transport protein to diffuse into or out of a cell?
a. $\mathrm{Ca}^{2+}$
b. $\mathrm{O}=\mathrm{C}=\mathrm{O}$
c. $\mathrm{O}=\mathrm{O}$
d. $\mathrm{H}-\mathrm{O}-\mathrm{H}$
6. A simple carbohydrate (e.g. glucose) needs to be brought into a muscle cell. Assuming the gradient is favorable, how will the molecule move in?
a. Simple diffusion, because carbohydrates are small and nonpolar.
b. Active transport, because the cell spends energy to move large molecules.
c. Facilitated diffusion, because carbohydrates are mid-sized, polar molecules.
d. Simple diffusion, because carbohydrates are small and polar.
e. Facilitated diffusion, because carbohydrates are nonpolar.

An experimenter divided 150 amoebas (a single-celled organism) between three petri dishes, then filled each dish with a specific salt solution to see how resilient amoebas are to fluctuating salt concentrations. The amoebas contain $25 \%$ salt.



50 amoebas in $30 \%$ salt solution


50 amoebas
in
$50 \%$ salt solution
7. Water will move $\qquad$ .
a. Into the amoebas in treatments $\mathrm{A}, \mathrm{B}$, and C.
b. Out of the amoebas in treatments $\mathrm{A}, \mathrm{B}$, and C .
c. Into amoebas in treatment A and out of amoebas in treatments B or C.
d. Into amoebas in treatment C and out of amoebas in treatments A or B .
e. Into amoebas in treatment B and out of amoebas in treatments A or C.
8. In the amoeba experiment above, lowering the salt concentration from $10 \%$ to $3 \%$ in treatment A would
a. Increase the rate of osmosis.
b. Stop osmosis.
c. Decrease the rate of osmosis.
d. Not affect the rate of osmosis.

## Photosynthesis and Cellular Respiration.

9. A mature maple tree can have a mass of 1 ton or more (dry biomass, after removing the water), yet it starts from a seed that weighs less than 1 gram. Which of the following processes contributes the most to this huge increase in biomass?
a. Absorption of mineral substances from the soil via the roots.
b. Absorption of organic substances from the soil via the roots.
c. Incorporation of $\mathrm{H}_{2} \mathrm{O}$ from the soil into molecules by green leaves.
d. Absorption of solar radiation into the leaf.

e. Incorporation of $\mathrm{CO}_{2}$ gas from the atmosphere into molecules by green leaves.
10. You eat a grape, which is high in glucose content. How could a glucose molecule from the grape provide energy to move your little finger?
a. The glucose is digested into $\mathrm{CO}_{2}$. and $\mathrm{H}_{2} \mathrm{O}$, which have more energy than the original glucose molecule and are used to fuel your finger cells.
b. The glucose molecule is rearranged into ATP molecules. The ATP fuels your finger cells.
c. The glucose molecule is energy and directly fuels your finger cells.
d. The energy of the glucose molecule is transferred to ATP as glucose is broken down. The ATP fuels your finger cells.
11. Which of the following best describes how a plant cell gets the energy it needs for cellular processes?
a. Chloroplasts make sugars, which are used in cellular respiration to make ATP.
b. Chloroplasts make ATP and then transport it to cells that lack chloroplasts (roots, stems, etc.).
c. In the light, ATP comes from chloroplasts, in the dark, from mitochondria.
d. Plants derive most of their ATP from organic matter absorbed by roots; the remaining ATP comes from the chloroplasts.
12. Review the figures depicting gas concentrations within a cell to determine which statement is true.
a. Figure 1 accurately represents gas concentrations during photosynthesis.
b. Figure 2 accurately represents gas concentrations during cellular respiration.
c. Figure 3 accurately represents gas concentrations during photosynthesis.
d. Figure 3 accurately represents gas concentrations for both cellular respiration and photosynthesis.




## Cell Division.

13. Answers (a) through (d) represent sperm cells (just focusing on chromosomes 1-4 and the sex chromosomes). Which is an accurate representation of a sperm cell? M stands for maternal chromosome and P stands for paternal chromosome.

14. Here is a karyotype of a human skin cell in $\mathrm{G}_{0}$ phase. This karyotype $\qquad$ .
a. should have fewer autosomes.
b. shouldn't have sex chromosomes.
c. should have a Y chromosome.
d. should have a second set of autosomes.

15. If dandelions have 16 total chromosomes, how many chromosomes are packaged into egg or pollen (the plant's gametic cells)?
a. 32
b. 25
c. 16
d. 8
16. If dandelions have 16 total chromosomes, how many chromosomes are packaged into an individual leaf cell?
a. 32
b. 25
c. 16
d. 8

Use the cell cycle figure to answer the next two questions.
17. DNA exists as chromosomes during part of this phase:
a. G1 phase
b. S phase
c. G 2 phase
d. Mitosis and cytokinesis
e. G0 phase

18. Recall the last time you cut your finger. The new cells that grew to replace the damaged cells were fully functional adult cells once they were in this phase:
a. G1 phase
b. S phase
c. G 2 phase
d. Mitosis and cytokinesis
e. G0 phase
19. Which statement is true?
a. Chromosomes are made up of nucleotides and proteins.
b. A DNA molecule is the condensed form of a chromosome.
c. DNA molecules are composed of chromosomes.
d. Proteins are composed of DNA molecules.
20. Genes $\qquad$ .
a. are made up of DNA molecules.
b. are made up of chromosomes.
c. are made up of nucleotides.
d. are made up of amino acids.

## Genetics.

Epistasis: Human skin color follows polygenic inheritance (A, B, and C genes). Dark skin is completely dominant to light skin. Albinism (complete lack of pigment) follows epistasis. Albinism is autosomal recessive (D gene).
21. Julie has the following genotypes that affect her skin color:

- Skin pigmentation genotype: AaBBcc
- Albinism genotype: dd.

Which statement is true regarding Julie's phenotype?
a. Julie has intermediate skin color.
b. Julie has dark skin color.
c. Julie has light skin color.
d. Julie is albino.
22. Refer to the question above. Which answer option would you have chosen if Julie's genotype for producing pigment was Dd?
a. Julie has intermediate skin color.
b. Julie has dark skin color.
c. Julie has light skin color.
d. Julie is albino.
23. There are people in Susan's family who have had Polycystic Kidney Disease (PKD). PKD is a single-gene disease in which clusters of fluid-filled sacs (cysts) form in the kidneys, often leading to kidney failure by the age of 10 and a reduced lifespan. Below is a list of facts that she has gathered from researching 5 generations of her family. Help her to draw the correct conclusion based on these facts.

- There is an equal probability of PKD affecting men and women.
- Symptoms seem to "disappear" in some generations.
- Her mother had genetic testing done and one gene showed PKD but she doesn't have any symptoms.
a. PKD is a sex linked disease
b. PKD is a recessive disease.
c. PKD is due to a single random mutation that is not heritable.
d. PKD is a polygenic; the more genes that are mutated, the sicker the individual is.

24. Sex-linked Condition: Darcy has noticed that her mom is showing signs of male pattern baldness (heritable, sex-linked on the X chromosome, recessive condition) and she is getting worried that she may have it to. What are the odds that Darcy will have male pattern baldness if her father Richard wasn't bald?
a. $0 \%$
b. $25 \%$
c. $50 \%$
d. $100 \%$

## Protein Synthesis.


25. Which molecule is not directly involved in translation?
a. DNA
b. mRNA
c. tRNA
d. rRNA
26. Individuals who have sickle cell disease (autosomal recessive disease) make red blood cells that are misshaped. Having two alleles with the sickle cell mutation results in errors to $\qquad$ .
a. Polypeptide chain folding
b. RNA sequence
c. The sequence of amino acids in the polypeptide chain
d. All of the above
27. Here is a section of genetic code for a healthy red blood cell.

First, determine if this code is DNA or mRNA.
Second, use the mRNA table to determine the healthy polypeptide sequence.

- UGU - CGA - CAC - AGU -


28. This is what the mutated, sickled red blood cell gene looks like:

- ACA - TCT - GTG - TCT-

Use the information in the question above to determine where the mutation(s) occurred. Note that one of the mutations was accommodated by amino acid redundancy. In which location of the code did the redundancy occur?


## Evolution.

29. Many infectious diseases are becoming difficult to treat because of bacterial resistance to antibiotics. Populations of bacteria can become resistant when they are exposed to an antibiotic. What is the best general explanation for how this occurs?
a. Over time, the antibiotic triggers the bacteria's immune system to destroy the antibiotic so that it can live.
b. The antibiotic activates enzymes within the bacteria cells, which destroys the antibiotic and allow the bacteria to live.
c. The antibiotic causes bacteria to mutate, so that resistant bacteria are more likely to arise.
d. The antibiotic kills all the bacteria that do not have antibiotic-resistant mutations. Resistant bacteria survive and reproduce.
30. A subset of butterflies from an ancestral population migrates to a new, isolated area where the species did not previously exist. The force of evolution that describes this is $\qquad$ and it leads to an immediate $\qquad$ of alleles within the ancestral population.
a. Gene flow ... increase
b. Genetic drift ... loss
c. Natural selection .... increase
d. Genetic drift ... increase
e. Gene flow ... loss
31. Choose the word or statement that best completes this sentence:
 Inbreeding $\qquad$ natural selection.
a. works in opposition to
b. stops
c. creates mutations that strengthen
d. encourages
32. Choose the word or statement that best completes this sentence: Gene flow $\qquad$ natural selection.
a. works in opposition to
b. stops
c. creates mutations that strengthen
d. encourages

## Ecology.

33. What immediate effect would a decrease in foxes have on this food web?
a. Buzzard populations would decrease.
b. Plantain populations would decrease.
c. Rabbit populations would decrease.
d. Mouse populations would decrease.


BIOLOGY 101 Fall, 2019
This figure depicts cougar populations in Washington State.

34. Use his figure to calculate the average population size of cougar from 1976-2011. How many deer would need to be present to support the average cougar population?
a. 10-100 deer
b. 1000 deer
c. $10,000-50,000$ deer
d. 500,000 deer
e. Millions of deer
35. In eutrophication, what is directly responsible for lowering $\mathrm{O}_{2}$ levels?
a. Nitrogen
b. Phosphorus
c. Phytoplanknton
d. Zooplankton
e. Bacteria
36. Consider the figure below depicting two species of paramecium grown in separate flasks (A) and together in the same flask (B). These two species do not consume one another. The results of figure (B) indicate that:

a. Competition occurred without niche partitioning
b. Competition occurred with niche partitioning
c. P. caudatum had adaptations that allowed it to survive in culture with $P$. Aurelia.
d. $\quad$. aurelia and $P$. caudatum engaged in a successful mutualism.

